Genetic susceptibility for inflammatory bowel diseases

Weersma, Rinse Karel

GENETIC SUSCEPTIBILITY FOR INFLAMMATORY BOWEL DISEASES

R.K.WEERSMA

The studies presented in this thesis were supported by the J.K. de Cock Foundation, Schering-Plough and the Dutch Society for Gastroenterology.

Financial support of the following sponsors for the printing of this thesis is gratefully acknowledged: Altana Pharma, Astra Zeneca, Boston Scientific, Bristoll-Myers Squibb, Ferring, Janssen-Cilag, Olympus, Schering-Plough, Tramedico, Vandeputte medical.

CIP GEGEVENS KONINKLIJKE BIBLIOTHEEK DEN HAAG

Weersma, Rinse Karel
Genetic susceptibility for inflammatory bowel diseases
Proefschrift Groningen.

ISBN: 978-90-367-3115-7

Layout: Helga de Graaf - Studio Eye Candy, Groningen (www.proefschrift.info)
Cover painting by Julie Mehretu
Printed by Printpartners Ipskamp, Enschede

RIJKSUNIVERSITEIT GRONINGEN

GENETIC SUSCEPTIBILITY FOR INFLAMMATORY BOWEL DISEASES

Proefschrift

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
woensdag 17 oktober 2007
om 16.15 uur

door

Rinse Karel Weersma

geboren op 5 oktober 1972

te Delfzijl

Promotores:	Prof. dr. J.H. Kleibeuker
	Prof. dr. C. Wijmenga
Copromotores:	Dr. G. Dijkstra
	Dr. H.M. van Dullemen
Beoordelingscommissie:	Prof. dr. D.W. Hommes
	Prof. dr. R.M.W. Hofstra
	Prof. dr. R.J. Ploeg

Paranimfen:	Drs. O.R. Maarsingh
	Dr. J.C. van Zanten

Contents

Chapter 1

Introduction and outline of the thesis

Introduction

Chronic inflammatory bowel diseases (IBD) comprising Crohn's disease (CD) and ulcerative colitis (UC) are characterized by chronic relapsing inflammation of the gastrointestinal tract. The combined prevalence of CD and UC is estimated at 100 – 200 / 100.000 in developed countries. The pathogenesis of IBD is only partially understood but concordance rates in twins and siblings suggest that a genetic predisposition, apart from environmental and immunological factors, contributes to the pathogenesis of IBD. In the past decade, tremendous progress has been achieved in unraveling the genetic etiology of IBD. By conducting genome wide scans, several susceptibility loci for IBD have been identified. In 2001 the *CARD15* gene encoding for the NOD2 protein on chromosome 16 (IBD1) has been found to be strongly associated with CD susceptibility.[1] NOD2 is part of the innate immune system and is an intracellular pathogen-associated molecular pattern (PAMP) receptor that recognizes specific bacterial membrane components. Two missense mutations (R702W and G908R) and one frameshift insertion mutation (L1007fsinsC) in the leucine rich repeat region of the protein are independently associated with ileal CD in Caucasian patients. The L1007fsinsC mutation causes a truncated protein, suggesting that a defect in bacterial recognition might be involved in CD. The exact mechanism how mutations in *CARD15* are involved in IBD susceptibility is still only partially understood and studies thus far give conflicting data. Although the discovery of *CARD15* as a susceptibility gene has not led to clinical consequences, important progress and insight in the pathogenesis of IBD has been made. Actually, the discovery of the *CARD15* gene in IBD has been one of the success stories in genetic research in complex genetic diseases and it has helped researchers to focus at signaling pathways of bacterial products in both epithelial and immune cells in the gut. Since CD and UC have many different phenotypic presentations and genome wide scans detected several linkage regions, IBD is considered a polygenic disease. Therefore many additional candidate-gene studies have been performed in the last decade.

CODE Study

For genetic research in complex genetic disorders as IBD it is important to have large homogenous cohorts of well described patients. In 2001 the CODE study (Chronische Ontsteking van de Darm en Erfelijkheid: Chronic Inflammation of the Gut and Inheritance) started in the University Medical Center Groningen. DNA has been, and is still being collected to form a large cohort of mainly Caucasian patients and family members from the Northern part of the Netherlands. Since the immigration rate in the Northern part of the Netherlands has been relatively low, this cohort can be considered a founder population. It is therefore suitable for genetic research and in particular for haplotype analysis, since it is assumed that the present population descends from a limited number of founders that results in evolutionary conserved haplotypes. Previously, several association studies have been performed in this cohort. For *CARD15* the R702W and the 1007fsinsC mutation were independently associated with CD and not with UC. For different subsets of CD, association was found for an early age of onset, ileal localization, familial occurrence of IBD and penetrating or stricturing disease behavior.[2] In another study in collaboration with the University Medical Center Nijmegen and

the Erasmus Medical Center Rotterdam association analysis was performed for IBD and Toll Like Receptor 4 (*TLR4*). TLR4 is, like NOD2, part of the innate immune system and is involved in NF-κB regulation. Haplotype analysis showed an association of CD and UC with *TLR4*, but in contrast with prior publications, no association could be found for the Asp266Gly and the Thr399Ile polymorphisms and IBD.[3] Furthermore a candidate gene study was carried out for association of IBD and the Multi Drug Resistant 1 (*MDR1*) Gene, but no association could be found with IBD, CD, UC or different subsets of patients.[4]

Genotypes vs. Phenotypes

An important aspect in studying IBD genetics is the consequent description of disease phenotypes. Since IBD is considered a multigenic disorder, different genes are probably involved in different subsets of phenotypes. It is therefore mandatory to have internationally accepted classification systems for IBD. An accepted and frequently used system is the Vienna classification.[5] It includes age of onset, disease localisation and disease behaviour. A number of studies, including a study from the CODE cohort has validated this classification. However, several considerations have led to an update of the Vienna classification system during an expert meeting in Montreal in 2005. The main modifications were the introduction of an early age of onset category (< 16 years), the possibility of co-classification of upper gastrointestinal involvement and the inclusion of perianal disease as a disease modifier instead of being a form of penetrating disease. For the current thesis the original Vienna classification has been used.

Pharmacogenetics

Pharmacogenetics is another research subject in IBD genetics. There has been much interest in the pharmacogenetics of azathioprine metabolism. Azathioprine is a purine analogue that is frequently used in the treatment of Crohn's disease but its use is hampered by the frequent occurrence of side-effects. Polymorphisms in the thiopurinemethyltransferase (*TPMT*) gene, which metabolizes azathioprine to 6-mercaptopurine and 6-methyl-mercaptopurine, and inosine triphosphate pyrophospatase (ITP-ase) deficiency which leads to accumulation of the metabolite 6-thio-ITP, have been found to be responsible for a subset of the side-effects of azathioprine therapy.

Aims and outline of the thesis

This thesis aims to gain insight in the genetic background of IBD.

The first part of this thesis focuses on specific genetic associations with IBD. Therefore, a detailed review of the current literature on IBD-genetics is given in **Chapter 2.** Since the discovery of the association of *CARD15* and IBD, many additional genes have been studied. Several of these genes are potentially truly associated, but results have been conflicting for many of the associations found. Next to specific genetic associations, current research on the functional role of mutations in *CARD15* in IBD is also reviewed.

The initial part of the thesis comprises two studies investigating novel candidate genes for IBD susceptibility. In **chapter 3** the association between IBD and Interleukin Receptor associated Kinase-M (*IRAK-M*) is studied. IRAK-M is a NF- κB -mediated, negative regulator of Toll-like receptor (TLR) signaling and is localized on chromosome 12q14, a susceptibility locus for IBD. It was hypothesized that a functional mutation in a negative regulator of TLR signaling might induce impaired endotoxin tolerance and increased inflammatory responses. Therefore *IRAK-M* is a good candidate gene for association analysis with IBD. 542 patients with IBD (309 CD and 233 UC) and 305 controls were studied. Phenotypic details of all CD patients according to the Vienna classification were available. UC patients were phenotyped according to an accepted classification including extend of the disease, age of onset, need for colectomy, extraintestinal manifestations and the occurrence of malignancy. Two single nucleotide polymorphisms (SNPs) and six microsattelite markers were evaluated by association analysis and Haplotype Sharing Statistics. Results were stratified for *CARD15* mutations R702W, G908R and 1007fsinsC.
In **Chapter 4** the genetic association between *RUNX3* and IBD is studied. *RUNX3* is a member of the runt domain family of transcription factors. It is known that loss of RUNX3 function is associated with a spontaneous colitis in knockout mice. It is a member of the TGF- ß signaling pathway, which is a potent inhibitor of inflammation in IBD. Impaired activation of RUNX3 might result in decreased activity of the TGF- ß pathway and decreased inhibition of inflammation in IBD.The gene encoding for *RUNX3* resides on chromosome 1p36, which is a susceptibility locus for IBD. Therefore *RUNX3* is a good candidate gene for susceptibility for IBD. Four SNPs and four microsattelite markers were studied for *RUNX3* in the CODE cohort. Furthermore, mutations in *SLC22A4* and *5* encoding for the organic cation transporters 1 and 2 (OCTN1/2) were found to be associated with CD in previous publications and an association was found between polymorphisms in *SLC22A4*, resulting in a disrupted binding site for RUNX in rheumatoid arthritis.[6] For that reason, association analysis for 6 SNPs in *SLC22A4/5* (including the known polymorphisms 207 G→C, 1672 C→G) and IBD and interaction with *RUNX3* was studied. All results were stratified for *CARD15* status. In addition to the genetic association analysis RUNX3 and OCTN1 expression was analyzed in colonic and ileal, inflamed and non-inflamed mucosal tissue samples of 30 IBD patients and 6 controls.

The second part of the thesis comprises three studies aimed at the confirmation of previously described genetic associations with IBD and describes specific genotype-phenotype interactions. Next to *CARD15*, *SLC22A4/5* and *TLR4,* several other genes have been identified to be associated with IBD susceptibility. Simultaneously with the identification of *SLC22A4/5*, genetic variations in *DLG5* (Drosophila Discs Large Homologue 5) on chromosome 10q23 showed association with CD.[7] DLG5 is important in maintaining epithelial stability and genetic variants could result in an impaired intestinal permeability. Additionally, two recent important studies identified two novel CD associated genes by performing the first genome wide association studies.[8 9] An uncommon coding SNP in the gene encoding for the interleukin-23 receptor (*IL23R*) conferred strong protection against CD. It was also shown to be associated with UC in non-Jewish patients. The other SNP in the autophagy-related 16-like 1 gene (*ATG16L1*) was shown to be associated with CD.

It is supposed that genetic susceptibility has a more prominent role in the aetiology of early-, than of late-onset IBD, since early-onset patients were less exposed to environmental factors than late-onset patients. As a result, a higher frequency of IBD associated mutations is expected. In **chapter 5** polymorphisms of *CARD15*, *TLR4*, *SLC22A4/5* and *DLG5* are analyzed in a cohort of 103 pediatric onset and 696 adult onset IBD patients and controls. Prevalence of mutations in the pediatric cohort was compared with the prevalence in adult-onset IBD and controls. Specific genotype-phenotype associations were studied.

Since it is of pivotal importance that genetic associations are confirmed in independent cohorts from different countries, **chapter 6** describes a replication study for the two most strongly associated SNPs in *IL23R* and *ATG16L1* in our cohort of IBD patients. We were also interested in discovering whether these two genes are more generally involved in other common chronic disorders of the gastrointestinal tract and we therefore included a cohort of celiac disease patients from the Netherlands.

It is not only mandatory that previously found associations are confirmed in independent cohorts, these cohorts also need to have sufficient power to detect specific genotype-phenotype interactions. For *SLC22A4/5* and *DLG5* there have been conflicting results in the literature, but many studies are hampered by small sample size or the lack of adequate uniform phenotypic descriptions. As mentioned before, for genetic research in complex genetic disorders as IBD it is important to have large homogenous cohorts of well described and uniformly phenotyped patients. For that reason a large nationwide collaborative project was initiated. Results are described in **chapter 7.** DNA samples and phenotypic details of IBD patients from seven University Medical Centers in the Netherlands (University Medical Center Groningen; Academic Medical Center, Amsterdam; VU University Medical Center, Amsterdam; Leiden University Medical Center; Radboud University Nijmegen Medical Center; Erasmus Medical Center, Rotterdam and the University Medical Center Utrecht) were collected. 2937 patients (1696 CD, 1099 UC and 142 with indeterminate colitis) and 1484 healthy controls were included. Phenotypic details were available for 2090 patients (1315 CD / 775 UC). We performed an association analysis for *DLG5, SLC22A4/5* and *ATG16L1* with IBD, CD, UC en different subsets of CD and UC. Interaction between these genes was studied.

The last part of the thesis focuses on azathioprine toxicity. In daily clinical practice it was noted that azathioprine toxicity occurred more often in IBD compared to other diseases for which azathioprine is frequently used. Particularly azathioprine induced pancreatitis had been rarely observed in other diseases than Crohn's disease in clinical practice as well as in the literature. To investigate this clinical observation, a retrospective case-note review has been performed analyzing azathioprine toxicity and necessity of withdrawal in 1564 patients with a liver or renal transplantation, systemic lupus erythematosus, Wegener's granulomatosis, autoimmune hepatitis, rheumatoid arthritis ulcerative colitis or Crohn's disease. Azathioprine use and toxicity in the University Medical Center Groningen were also compared to the use in IBD patients in a large community hospital (Martini Hospital Groningen). Results are described in **chapter 8**. The fact that azathioprine induced acute pancreatitis was more prevalent in Crohn's disease compared to other diseases, led to the exploration of an association of azathioprine induced pancreatitis and circulating pancreatic antibodies, which are highly specific for CD compared to UC and other autoimmune diseases.

Finally the results of the studies in this thesis are summarized and future perspectives for genetic research in IBD are given in **chapter 9**.

References

1. Hugot JP, Chamaillard M, Zouali H *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature. 2001;411:599-603.

2. Oostenbrug LE, Nolte IM, Oosterom E *et al*. *CARD15* in inflammatory bowel disease and Crohn's disease phenotypes: an association study and pooled analysis. Dig Liver Dis. 2006;38:834-45.

3. Oostenbrug LE, Drenth JP, de Jong DJ *et al*. Association between Toll-like receptor 4 and inflammatory bowel disease. Inflamm Bowel Dis. 2005;11:567-75.

4. Oostenbrug LE, Dijkstra G, Nolte IM *et al*. Absence of association between the multidrug resistance (MDR1) gene and inflammatory bowel disease. Scand J Gastroenterol. 2006;41:1174-82.

5. Gasche C, Scholmerich J, Brynskov J *et al*. A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. Inflamm Bowel Dis. 2000;6:8-15.

6. Peltekova VD, Wintle RF, Rubin LA *et al*. Functional variants of OCTN cation transporter genes are associated with Crohn's disease.Nat Genet. 2004;36:471-5.

7. Stoll M, Corneliussen B, Costello CM *et al*. Genetic variation in DLG5 is associated with inflammatory bowel disease. Nat Genet. 2004;36:476-80.

8. Duerr RH, Taylor KD, Brant SR *et al*. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science. 2006;314:1461-3.

9. Hampe J, Franke A, Rosenstiel P *et al*. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet. 2007;39:207-211.

Chapter 2

Inflammatory bowel diseases and genetics: current affairs

R.K. Weersma, H.M. van Dullemen, G. van der Steege,
I.M. Nolte, J.H. Kleibeuker, G. Dijkstra

Alimentary Pharmacology and Therapeutics, accepted for publication

Abstract

Tremendous progress has been achieved in unraveling the genetic etiology of inflammatory bowel diseases (IBD) comprising ulcerative colitis (UC) and Crohn's disease (CD). It has led to the discovery of mutations in NOD2/*CARD15* associated with ileal CD. It is only partially understood how mutations in NOD2/*CARD15* lead to CD. Mouse models, in vitro data and studies in humans offer conflicting data whether there is a loss or gain of function of NOD2 in CD. Through the conductance of genome wide scans and hypothesis driven candidate gene studies several additional genes have been identified. Only few of these genes are currently being recognized as potential disease causing or disease modifying genes. Promising candidate genes include Toll like receptor 4 (*TLR4*), Multi Drug Resistance 1 (*MDR1*), NOD1 (*CARD4*), HLA DRB1*103, *DLG5* as well as the IBD5 locus including members of the organic cation transporter cluster 1 and 2. For future genetic research accurate phenotyping of patients is very important and large population based cohorts are needed. Although genetic research has not yet led to better prediction of the disease course, development of malignancy or patient selection for medical therapy, remarkable progress has been made in the understanding of the pathogenesis of IBD. Eventually, genetic research may be able to classify different disease phenotypes on a more detailed molecular basis and may provide important contributions in the development of new therapeutic approaches.

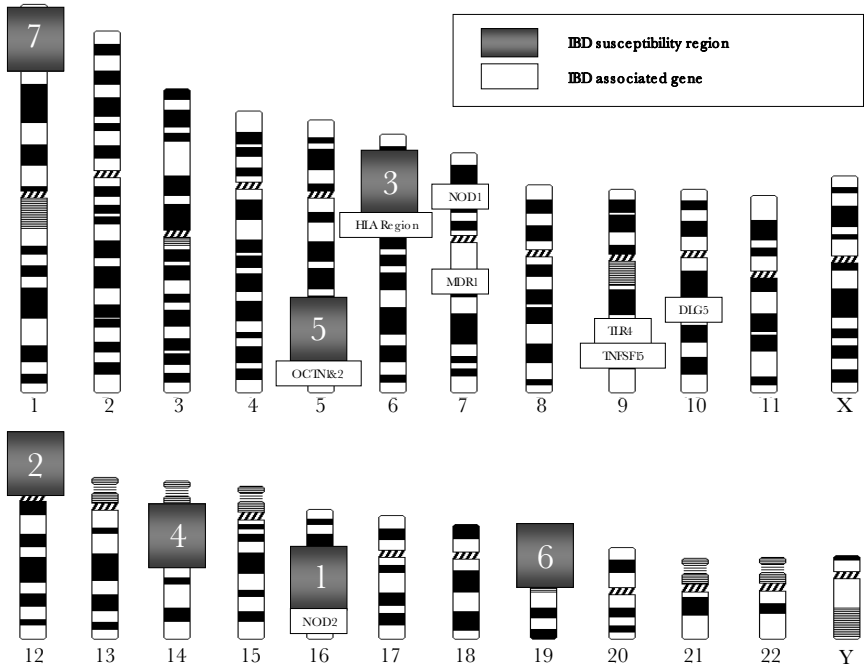


Figure 1
Inflammatory Bowel Diseases (IBD) associated genes and susceptibility loci for IBD on the human genome, termed IBD1 to IBD7.

Introduction

In the past decade, tremendous progress has been achieved in unraveling the genetic etiology of inflammatory bowel diseases (IBD) comprising ulcerative colitis (UC) and Crohn's disease (CD). By conducting genome linkage studies, several susceptibility loci for IBD have been identified termed IBD1 to IBD7 (figure 1).[1-12] Through fine mapping of these susceptibility loci as well as by hypothesis driven candidate gene studies, the *CARD15* gene encoding for the NOD2 protein on IBD1 has been found to be strongly related to CD susceptibility.[13 14] Since IBD is considered a multigenic disorder, many other candidate genes have been studied in addition to *CARD15*. None of these genes has been as consistently replicated as *CARD15* and probably other disease susceptibility genes will not be discovered as straightforward as *CARD15*, due to low penetrance and complex gene-gene and gene-environment interaction. Promissing candidate genes include Toll like receptor 4 (*TLR4*), Multi Drug Resistance 1 (*MDR1*), NOD1 (*CARD4*), HLA DRB1*103, *DLG5* as well as the IBD5 locus including members of the organic cation transporter cluster 1 and 2.[15-21]

In addition to the search for new IBD susceptibility genes, one of the main research themes in IBD focuses on the function of NOD2 in CD. Although many questions have been answered, it still remains unknown whether there is loss or gain of protein function in patients carrying mutated NOD2.[22]
We will summarize current topics in IBD genetics, focussing on NOD2 functionality and novel candidate genes.

Innate immunity

Since the discovery of NOD2 in CD susceptibility, the innate immune system has been studied extensively. Current research, regarding the innate immune system in IBD, is focussing on NOD2 functionality and new candidate genes for IBD susceptibility.

Function of NOD2

NOD2 is composed of two N-terminal caspase recruitment domains (CARDs), a nucleotide binding and oligomerisation domain and ten C-terminal leucine rich repeats (LRRs) (figure 2). It is mainly expressed in macrophages, neutrophils and dendritic cells as well as in intraepithelial Paneth cells that are located in the crypts of Lieberkühn in the ileum.[23,24] NOD2 is part of the innate immune system and is a pathogen associated molecular pattern (PAMP) receptor that recognizes muramyl dipeptide (MDP) which is a part of bacterial peptidoglycanes. The three variants associated with IBD are all located in the LRR region which is the binding site for MDP. They consist of two missense mutations (R702W and G908R) and one frameshift insertion mutation (L1007fsinsC).[25,26] The L1007fsinsC mutation causes a truncated protein, suggesting that a defect in bacterial recognition might be involved in CD. Interestingly, a mutation in NOD2 located in the central nucleotide binding domain instead of the LRR region is associated with Blau's syndrome, a rare autosomal dominant disorder with granulomatous arthritis, uveitis and skin rash, implying different pathological pathways.[27] CD is characterized by an increased activity of NF-κB and NOD2 has been shown to have a role in the activation of NF-κB. However, the precise mechanism, how mutations in NOD2 and subsequently activation of NF-κB leads to susceptibility of IBD is still only partially understood.[28] Studies in NOD2 deficient mice have given contradictory results and one seemingly simple question whether there is loss or gain of function in CD associated NOD2 variants remains to be answered.[29-31]
In a recent study with insertion of mutated NOD2 alleles in mice, Maeda *et al.* showed that mutant mice had increased activation of the NF- κB pathway, increased secretion of IL-1β and were more susceptible to bacteria-mediated experimental colitis, suggesting a gain of function.[29] On the other hand, a study by Watanabe and co-workers showed that NOD2 functions as a negative regulator of TLR2-mediated T helper type 1 response. NOD2 deficient mice showed increased TLR2-driven activation of NF- κB, particularly of the NF- κB subunit c-Re1, whereas intact NOD2 signaling inhibited TLR2-driven activation of NF- κB.[30] In an additional study, NOD2 deficient mice had decreased expression in intestinal Panethcells of cryptdins, which are mouse homologues of human α-defensins.These α-defensins are important mem-

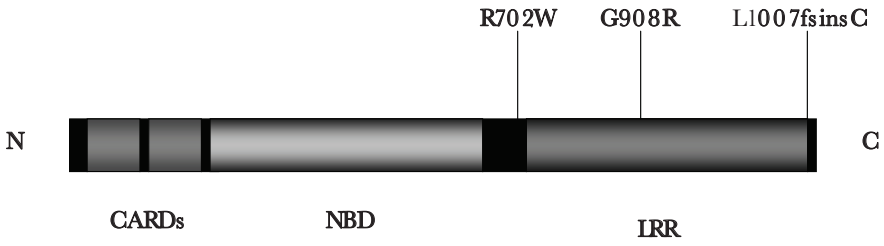


Figure 2
Structure of the NOD2 protein with the three variants in the LRR region associated with CD. Abbreviations: CARDs, caspase activating recruitment domains; NBD, nucleotide binding domain; LRR, leucine rich region.

bers of the innate immune system. NOD2 deficient mice were susceptible to bacterial infection via the oral route but not through intravenous or intraperitoneal exposure, suggesting an important role for NOD2 as a regulator of bacterial immunity in the intestine.[31] Indeed, reduced expression of α-defensins is observed in Paneth cells in subjects with CD, which is even more pronounced in patients with NOD2 mutations.[32 33] Loss of function of NOD2 may lead to enhanced mucosal invasion of bacteria resulting in an increased inflammatory response.
Still, mouse models, in vitro data and studies in humans offer conflicting data and many additional studies are currently being undertaken to give an answer to the question how mutations in the LRR region of NOD2 lead to CD.

NOD1 and Toll Like Receptors

Since the identification of NOD2 in CD susceptibility, many further candidate gene studies were performed in genes involved in the innate immunity. The most promising results, although data have been conflicting and not solidly reproduced, suggest an association of *NOD1/CARD4* and *TLR4* with IBD.[15 17]
NOD1 is expressed in the epithelium of the small and large intestine and is fairly similar to NOD2, but differs by the presence of only one CARD. Its LRR functions as a pattern recognition receptor for diaminopimelic acid present in gram-negative bacterial peptidoglycans. Activation of NOD1 activates NF- κB and enhances apoptosis.[34]
In addition, the gene encoding for NOD1 (*CARD4*) is located on chromosome 7p14, previously recognized as an IBD susceptibility locus.[10] Though a previous study did not show an association with IBD, NOD1 has recently been implied in IBD susceptibility.[17 35] McGovern *et al.* found that the deletion allele of a complex functional insertion-deletion polymorphism was associated with early age of onset IBD and extraintestinal manifestations.[17] The same study showed that haplotypes in the terminal exons of *CARD4* were also significantly associated with IBD. Again, these data need further confirmation.

The Toll Like Receptors are important members of the innate immunity. Different TLRs recognize selectively PAMPs of different classes of microorganisms.[36] TLR4 recognizes lipopolysac-

charidases which are components of the cellular wall of gram negative bacteria. TLR4 is up-regulated during intestinal inflammation in macrophages, dendritic cells and epithelial cells in IBD.[37] Upon binding with LPS it forms a complex with CD14 at the surface of monocytes and neutrophils leading to NF- κB activation and the release of inflammatory cytokines.[38 39] The Asp299Gly mutation is located in the extracellular LRR domain of TLR4 and is associated with decreased responsiveness to endotoxins in humans.[40] This mutation has been found to be associated with CD in two studies and with IBD in one Belgian study.[41-43] A second mutation Thr399Ile which is in strong LD with Asp299Gly has also been implied with IBD susceptibility.[44] These results have not been confirmed in several other studies.[15 45] In a study by our own group, association with the *TLR4* region was found by haplotype analysis of microsatellite markers surrounding the gene but not with Thr399Ile or Asp299Gly.[15] This finding implies that the Thr399Ile and Asp299Gly polymorphisms could be merely in LD with other disease associated variants in the *TLR4* region, instead of being causative themselves.
A recent study found no genetic association of TLR1, TLR2 and TLR6 with disease susceptibility but an association between variants of these TLRs and a specific phenotype of extensive colonic disease in CD or UC was established.[46]

The HLA region

The HLA region on the short arm of chromosome 6p (IBD3) was identified as a disease susceptibility locus in several genome wide scans as well as in a recent meta-analysis of 10 genome wide scans [4 6 9 12 47] This area harbors a total of 224 highly polymorphic genes, many of which appear to have immunoregulatory functions.[48] HLA proteins present peptides to T-Cell receptors and are divided in class I and II HLA proteins. The HLA class II proteins play a central role in the immune response, are expressed on specialized immune cells and consist of an α-chain and a ß-chain that form a groove by which the antigen peptide is presented to the T cell receptor.[49] The three genes encoding for these α- and ß-chains are HLA-DP, HLA-DQ and HLA-DR which are highly polymorphic. HLA class II genes have been extensively studied in association studies for IBD.[50]
A meta-analysis of data published until 1998 showed a positive association with UC of HLA DR2, DR9 and DRB*103 and a protective association with DR4. For CD, HLA-DR7, DRB3*0301 and DQ4 showed a positive association whereas HLA DR2 and DR4 showed a protective association.[52] The association with HLA- DRB*103 has been confirmed in European cohorts and seems to be associated with extensive disease in UC and in CD with pure colonic involvement. HLA- DRB*103 seems to be a potentially important contributor to colonic involvement of IBD. However, the rare variant of HLA- DRB*103 has a very low frequency in the general population thereby minimizing its possible clinical value.[53 54] Polymorphisms for HLA-DRB1*1502 which are found in populations with variable ethnic background have been associated with UC.[57-59] HLA DRB*07 has been consistently replicated in CD patients with ileal localisation. Interestingly, patients carrying the risk associated allele for HLA DRB*07 did not carry any of the three NOD2 mutations, stressing the fact that there is a great genetic heterogeneity in CD,

even within specific phenotypes such as ileal localisation.[55 56]
Tumour Necrosis Factor-α (TNFα) is a proinflammatory cytokine that plays an important role in the inflammatory processes involved in IBD.[60 61] The importance of TNFα is highlighted by clinical efficacy of treatment with anti-TNFα monoclonal antibodies in CD.[62 63] TNF polymorphisms have been extensively studied in the HLA class III region. Several polymorphisms in the promoter region are known, but functional data are lacking and genetic association studies have been very inconsistent.[64 65] Another interesting finding, although not in the HLA region, is the recently found association with the Tumour Necrosis Factor superfamily member 15 (*TNFSF15*) on chromosome 9q33 in a Japanese and a UK cohort. The function of *TNFSF15* remains to be elucidated but it is involved in NF-κB activation and induction of apoptosis.[66 67]
In conclusion, several genes in the HLA complex are involved in IBD susceptibility, but results of genetic research have been conflicting because of the high density of genes in this region, a high degree of LD resulting in highly conserved haplotypes and complex gene-gene interactions with genes outside the HLA region.[51]

Novel candidate genes

IBD5

A genome wide screen in affected Canadian families identified linkage with a region at chromosome 5q31-q33 (IBD5) with a length of approximately 250 kb conferring susceptibility for CD with early age of onset.[9 20] This region contains a number of potentially interesting genes, including several genes encoding for immunoregulatory cytokines. However, identifying the causative gene(s) in the IBD5 locus has been hampered by the high extent of LD in the region. Recently, two functionally relevant mutations in the carnitine / organic cation transporter (OCTN) genes on the IBD5 locus were shown to be associated with CD.[21] The T substitution of C1672 in exon 9 of the *SLC22A4* gene encoding for OCTN1 and a G to C substitution at position -207 in the promoter region of *SLC22A5* gene encoding for OCTN2, were identified. Together these polymorphisms form a 2 allele risk haplotype associated with CD susceptibility. Gene-gene interaction with mutations in the *CARD15* gene is suggested.[68]
These results have been replicated in a German population even though the association was much weaker than in the original study of Peltekova *et al.*[69] This is in concordance with the findings that the IBD5 locus as a whole is less important in European populations compared to the original Canadian cohort.[70] Other studies did not find an association between OCTN polymorphisms and CD, though in one study an increased risk in a subgroup with perianal Crohn's disease has been found.[71 72]
Relatively little attention has been given to these genes before, due to the lack of a sensible explanation for a role in the pathogenesis of IBD. In the paper by Peltekova *et al.* preliminary functional experiments were performed demonstrating that the two single nucleotide polymorphisms (SNPs) impair function of the OCTNs, resulting in reduced carnitine transport, but convincing functional data remain scarce.

In addition to the lack of solid replication studies and functional data, it is also difficult to understand how two mutations in two adjacent genes produce IBD susceptibility. It still remains to be seen whether the *SLC22A4* and *SLC22A5* genes encoding for OCTN1 and OCTN2 are the causative genes for IBD susceptibility at the IBD5 locus or that they are merely haplotype tagging SNPs in strong linkage disequilibrium with other disease associated genes. Larger population based cohorts and functional data are needed to answer this important question.

DLG5

Recently the *DLG5* gene (Drosophila Discs Large Homologue 5) has been recognized as a novel susceptibility gene for CD and IBD.[19] *DLG5* is important in maintaining the epithelial structure and genetic variants could result in impaired intestinal permeability.[73]

Stoll *et al* identified *DLG5* by refining their previously defined linkage region on chromosome 10q23. It is worth mentioning that this region has not been confirmed in other genome wide scans. Genetic variants in *DLG5* were identified by using a positional cloning strategy. Two haplotypes were involved in IBD and CD susceptibility. A SNP 113 G→A, resulting in an aminoacid substitution R30Q, was positively associated with IBD and CD patients in a case control study. A second haplotype, identified by several other tagging SNPs was protective for IBD. Gene-gene interaction with *CARD15* in CD was detected by a significant difference in association of the 113A variant in patients carrying the risk alleles for *CARD15* compared to patients not carrying these alleles. Only one published study has been able to reproduce these results. [74] Initial enthusiasm has been discouraged by several studies from Europe and Japan, failing to show any association for *DLG5* with IBD. [75-78] In the same way as for OCTN, large population based case-control studies are needed to elucidate the role of *DLG5* in IBD susceptibility.

MDR1

P- glycoprotein-170 (Pgp170) is encoded by the multi drug resistance gene (*MDR1*). It was initially recognized to be responsible for resistance to cytotoxic drugs in cancer cells. It is an ATP binding cassette (ABC) transporter which is highly expressed in the intestinal epithelium. There is mounting evidence that Pgp170 is an important factor in host-bacterial interactions and in maintaining intestinal homeostasis.[79] Mice lacking the *MDR1* gene spontaneously develop colitis which is histologically similar to UC.[80] In combination with the fact that *MDR1* is located on chromosome 7q22 which was identified as an IBD susceptibility locus in one UK cohort *MDR1* appears to be a good candidate gene.[10]

Two exonic SNPs C3435Tand G2677T/A are involved in altered activity and expression of PgP170. Schwab *et al.* found the T allele and TT genotype of the C3453T SNP to be associated with UC and not with CD.[16] This association with UC has been confirmed in a Slovenian and Scottish cohort and with IBD in a mixed North American population.[81-83] This is in contrast with two other studies that could not replicate these positive findings.[84 85] In a large case-control study of Dutch patients in our own center, we did not find an association with the *MDR1* gene and IBD, UC or CD, neither on a single locus and haplotype association analysis nor with the haplotype sharing statistics.[86]

Although the *MDR1* gene holds promise as being involved in IBD susceptibility, there is doubt whether the previously found associations are true positive results.

Conclusions and future perspectives

Research in the genetic background of IBD is a rapidly evolving field and has helped unravelling pathophysiological processes involved in complex multifactorial diseases as IBD. Important progress and insight in the pathogenesis of IBD has been made by the discovery of NOD2/*CARD15*. In fact, the discovery of the *CARD15* gene in IBD has been one of the success stories in genetic research in complex genetic diseases. The exact mechanism how NOD2 is involved in IBD susceptibility is unknown to date, but genetic research has helped us to focus research at signaling of bacterial products in both epithelial and immune cells in the gut.
Furthermore, numerous candidate gene studies have been performed which has resulted in several positive and negative associations for IBD. Many of these studies are hampered by small sample sizes and hence the lack of power to detect low penetrant genes or complex gene-gene interactions. Only few of the identified genes are currently being recognized as potential disease causing or disease modifying genes. To establish true positive associations, large population based case-control studies and functional studies are needed in the future. It is therefore encouraging that collaboration between centers and countries is initiated in IBD genetics.
In addition to susceptibility for IBD, genetic variations may account for disease expression including disease location, clinical behaviour and response to therapy. For that reason, one important "current affair" in IBD genetics is the reliable phenotyping of patients. Since the successful connection of NOD2 with ileal CD and the increasing number of serological markers associated with different subsets of IBD, it is recognized that different phenotypes of CD or UC patients are characterized by different molecular and serological markers.[87 88] For future research, it is therefore of utmost importance that patients are accurately phenotyped according to a well defined clinical classification scheme. The Vienna classification is frequently used for genetic studies in CD and includes age of onset (A), disease localisation (L) and diseases behaviour (B).[89] A number of studies have validated this classification; however, several considerations have led to an update of the Vienna classification system during an expert meeting in Montreal in 2005.[90 91] The main modifications were the introduction of an early age of onset category (< 16 years), the possibility of co-classification of upper gastrointestinal involvement and the inclusion of perianal disease as a disease modifier instead of being a form of penetrating disease. The latter is an important modification because it is recognized that there is no clear association between perianal disease and intra-abdominal penetrating disease.[92] The aim of rigorous phenotyping is to include serological and genetic markers into the clinical classification system to stratify patients and eventually to predict disease course and response to medical therapy.
Although genetic research has not yet led to better prediction of the disease course, development of malignancy or patient selection for medical therapy, remarkable progress has been made in the understanding of the pathogenesis of IBD. Genetic research holds a strong promise for the future of IBD research. Eventually, genetic research may be able to classify different disease phenotypes on a more detailed molecular basis and may provide important contributions in the development of new therapeutic approaches.

References

1. Hugot JP, Laurent-Puig P, Gower-Rousseau C *et al*. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature 1996;379:821-3.

2. Cho JH, Nicolae DL, Gold LH *et al*. Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: evidence for epistasis between 1p and IBD1. Proc Natl Acad Sci U S A 1998;95:7502-7

3. Duerr RH, Barmada MM Zhang L, Pfutzer R, Weeks DE. High-density genome scan in Crohn disease shows confirmed linkage to chromosome 14q11-12. Am J Hum Genet 2000;66:1857-62.

4. Hampe J, Schreiber S, Shaw SH *et al*. A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort.Am J Hum Genet. 1999;64:808-16.

5. Ma Y, Ohmen JD, Li Z, Bentley LG. A genome-wide search identifies potential new susceptibility loci for Crohn's Disease. Inflamm Bowel Dis. 1999;5:271-8.

6. Williams CN, Kocher K, Lander ES *et al*. Using a genome-wide scan and meta-analysis to identify a novel IBD locus and confirm previously identified IBD loci.Inflamm Bowel Dis. 2002;8:375-81.

7. Paavola-Sakki P, Ollikainen V, Helio T *et al*. Genome-wide search in Finnish families with inflammatory bowel disease providesevidence for novel susceptibility loci. Eur J Hum Genet. 2003;11:112-20.

8. Vermeire S, Rutgeerts P, Van Steen K *et al*. Genome wide scan in a Flemish inflammatory bowel disease population: support for the IBD4 locus, population heterogeneity, and epistasis. Gut. 2004;53:980-6.

9 . Rioux JD, Silverberg MS, Daly MJ *et al*. Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. Am J Hum Genet 2000;66:1863-70.

10. Satsangi J, Parkes M, Louis E *et al*. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nat Genet 1996;14:199-202.

11 . Rioux JD, Daly MJ, Silverberg MS *et al*. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genet 2001;29:223-8.

12. Barmada MM, Brant SR, Nicolae DL A *et al*. Genome scan in 260 inflammatory bowel disease-affected relative pairs. Inflamm Bowel Dis. 2004;10:15-22

13. Ogura Y, Bonen DK, Inohara N *et al*. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature. 2001;411:603-6.

14. Hugot JP, Chamaillard M, Zouali H *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature. 2001;411:599-603

15. Oostenbrug LE, Drenth JP, de Jong DJ *et al*. Association between Toll-like receptor 4 and inflammatory bowel disease. Inflamm Bowel Dis. 2005 ;11:567-75

16. Schwab M, Schaeffeler E, Marx C *et al*. Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. Gastroenterology. 2003;124:26-33

17. McGovern DP, Hysi P, Ahmad T *et al*. Association between a complex insertion/deletion polymorphism in NOD1 (CARD4) and susceptibility to inflammatory bowel disease. Hum Mol Genet. 2005;14:1245-50.

18. De la Concha EG, Fernandez-Arquero M, Lopez-Nava G *et al*. Susceptibility to severe ulcerative colitis is associated with polymorphism in the central MHC gene IKBL. Gastroenterology. 2000;119:1491-5.

19. Stoll M, Corneliussen B, Costello CM *et al*. Genetic variation in DLG5 is associated with inflammatory bowel disease. Nature genetics 2004;36:476-80.

20. Rioux JD, Daly MJ, Silverberg MS *et al*. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genet. 2001;29:223-8.

21. Peltekova VD, Wintle RF, Rubin LA, *et al*. Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nature genetics 2004;36:471-75.

22. Eckmann L, Karin M. NOD2 and Crohn's disease: loss or gain of function? Immunity. 2005;22:661-7.

23. Ogura Y, Inohara N, Benito A *et al*. Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB. J Biol Chem. 2001;276:4812-8

24. Ogura Y, Lala S, Xin W *et al*. Expression of NOD2 in Paneth cells: a possible link to Crohn's ileitis. Gut. 2003;52:1591-7

25. Ahmad T, Armuzzi A, Bunce M *et al*. The molecular classification of the clinical manifestations of Crohn's disease. Gastroenterology. 2002;122:854-66

26. Economou M, Trikalinos TA, Loizou KT *et al*. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. Am J Gastroenterol. 2004;99:2393-404

27. Miceli-Richard C, Lesage S, Rybojad M *et al*. CARD15 mutations in Blau syndrome Nat Genet. 2001;29:19-20.

28 . Ogura Y, Inohara N, Benito A *et al*. Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB. The Journal of biological chemistry 2001;276:4812-18.

29 . Maeda S, Hsu LC, Liu H *et al*. Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing. Science. 2005;307:734-8.

30. Watanabe T, Kitani A, Murray PJ *et al*. NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses. Nat Immunol 2004;5:800-808

31. Kobayashi KS, Chamaillard M, Ogura Y *et al*. Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. Science. 2005;307:731-4.

32. Wehkamp J, Salzman NH, Porter E *et al*. Reduced Paneth cell alpha-defensins in ileal Crohn's disease. Proc Natl Acad Sci USA. 2005;102:18129-34.

33. Wehkamp J, Harder J, Weichenthal M *et al*. NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression. Gut. 2004;53:1658-64.

34. Chamaillard M, Hashimoto M, Horie Y *et al*. An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. Nat Immunol. 2003;4:702-7

35. Zouali H, Lesage S, Merlin F *et al*. CARD4/NOD1 is not involved in inflammatory bowel disease. Gut 2003;52:71-4.

36. Cario E. Bacterial interactions with cells of the intestinal mucosa: Toll-like receptors and NOD2. Gut. 2005 ;54:1182-93

37. Hausmann M, Kiessling S, Mestermann S, *et al*. Toll-like receptors 2 and 4 are up-regulated during intestinal inflammation.Gastroenterology. 2002;122:1987-2000.

38. Poltorak A, He X, Smirnova I, Liu MY*et al*. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. Science 1998;282:2085-8.

39. Chow JC, Young DW, Golenbock DT Christ WJ, Gusovsky F. Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. J Biol Chem 1999;274:10689-92

40. Arbour NC, Lorenz E, Schutte BC *et al*.TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. Nat Genet 2000;25:187-91.

41. Gazouli M, Mantzaris G, Kotsinas A *et al*. Association between polymorphisms in the Toll-like receptor 4, CD14, and CARD15/NOD2 and inflammatory bowel disease in the Greek population. World J Gastroenterol 2005;11:681-5.

42. Ouburg S, Mallant-Hent R, Crusius JB *et al*. The toll-like receptor 4 (TLR4) Asp299Gly polymorphism is associated with colonic localisation of Crohn's disease without a major role for the Saccharomyces cerevisiae mannan-LBP-CD14-TLR4 pathway. Gut. 2005;54:439-40.

43. Franchimont D, Vermeire S, El HH, Pierik M *et al*. Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. Gut 2004;53:987-92.

44. Török HP, Glas J, Tonenchi L. *et al*. Polymorphisms of the lipopolysaccharide-signaling complex in inflammatory bowel disease: association of a mutation in the Toll-like receptor 4 gene with ulcerative colitis. Clin Immunol 2004 ;112:85-91.

45. Lakatos PL, Lakatos L, Szalay *et al*. Toll-like receptor 4 and NOD2/CARD15 mutations in Hungarian patients with Crohn's disease: phenotype-genotype correlations.World J Gastroenterol 2005;11:1489-95

46. Pierik M, Joossens S, Van Steen K. *et al*. Toll-like receptor-1, -2, and -6 polymorphisms influence disease extension in inflammatory bowel diseases. Inflamm Bowel Dis. 2006;12:1-8.

47. van Heel DA, Fisher SA, Kirby A *et al*. Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. Hum Mol Genet 2004;13:763-70

48. The MHC sequencing consortium. Complete sequence and gene map of a human major histocompatibility complex. Nature. 1999;401:921-3.

49. Jardetzky TS, Brown JH, Gorga JC *et al*. Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen. Nature. 1994 ;368:711-8

50. Yap LM, Ahmad T, Jewell DP. The contribution of HLA genes to IBD susceptibility and phenotype. Best Pract Res Clin Gastroenterol. 2004 Jun;18(3):577-96.

51. Ahmad T, Neville M, Marshall SE *et al*. Haplotype-specific linkage disequilibrium patterns define the genetic topography of the human MHC. Hum Mol Genet. 2003;12:647-56.

52. Stokkers PC, Reitsma PH, Tytgat GN, van Deventer SJ. HLA-DR and -DQ phenotypes in inflammatory bowel disease: a meta-analysis.Gut. 1999;45:395-401.

53. de la Concha EG, Fernandez-Arquero M, Lopez-Nava G *et al*. Susceptibility to severe ulcerative colitis is associated with polymorphism in the central MHC gene IKBL. Gastroenterology. 2000;119:1491-5

54. Silverberg MS, Mirea L, Bull SB *et al*. A population- and family-based study of Canadian families reveals association of HLA DRB1*0103 with colonic involvement in inflammatory bowel disease. Inflamm Bowel Dis. 2003;9:1-9

55. Newman B, Silverberg MS, Gu X, *et al*. CARD15 and HLA DRB1 alleles influence susceptibility and disease localization in Crohn's disease. Am J Gastroenterol. 2004;99:306-15.

56. Fernandez L, Mendoza JL, Martinez A *et al*. IBD1 and IBD3 determine location of Crohn's disease in the Spanish population. Inflamm Bowel Dis. 2004;10:715-22.

57. Myung SJ, Yang SK, Jung HY,*et al*. HLA-DRB1*1502 confers susceptibility to ulcerative colitis, but is negatively associated with its intractability: a Korean study. Int J Colorectal Dis. 2002;17:233-7.

58. Yoshitake S, Kimura A, Okada M Yao T, Sasazuki T. HLA class II alleles in Japanese patients with inflammatory bowel disease.Tissue Antigens. 1999;533:50-8.

59. Trachtenberg EA, Yang H, Hayes E HLA *et al*. Class II haplotype associations with inflammatory bowel disease in Jewish(Ashkenazi) and non-Jewish caucasian populations. Hum Immunol. 2000;61:326-33.

60. Neurath MF, Fuss I, Pasparakis M *et al*. Predominant pathogenic role of tumor necrosis factor in experimental colitis in mice. Eur J Immunol. 1997;27:1743-50.

61. Noguchi M, Hiwatashi N, Liu Z *et al*. Secretion imbalance between tumour necrosis factor and its

inhibitor in inflammatory bowel disease.Gut. 1998;43:203-9.

62. Van Dullemen HM, van Deventer SJ, Hommes DW, Bijl HA, Jansen J, Tytgat GN,Woody J. Treatment of Crohn's disease with anti-tumor necrosis factor chimeric monoclonal antibody (cA2). Gastroenterology. 1995;109:129-35.

63. Targan SR, Hanauer SB, van Deventer SJ *et al*. A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factoralpha for Crohn's disease. Crohn's Disease cA2 Study Group. N Engl J Med. 1997;337:1029-35.

64. O'Callaghan NJ, Adams KE, van Heel DA, Cavanaugh JA. Association of TNF-alpha-857C with inflammatory bowel disease in the Australian population. Scand J Gastroenterol. 2003;38:533-4.

65. Negoro K, Kinouchi Y, Hiwatashi N *et al*. Crohn's disease is associated with novel polymorphisms in the 5'-flanking region of the tumor necrosis factor gene. Gastroenterology. 1999;117:1062-8.

66. Yamazaki K, McGovern D, Ragoussis J *et al*. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. Hum Mol Genet. 2005;14:3499-506.

67. Migone TS, Zhang J, Luo X *et al*.TL1A is a TNF-like ligand for DR3 and TR6/DcR3 and functions as a T cell costimulator.Immunity. 2002;16:479-92

68. Newman B, Gu X, Wintle R *et al*. A risk haplotype in the Solute Carrier Family 22A4/22A5 gene cluster influences phenotypic expression of Crohn's disease. Gastroenterology. 2005;128:260-9.

69. Torok HP, Glas J, Tonenchi L *et al*. Polymorphisms in the DLG5 and OCTN cation transporter genes in Crohn's disease. Gut 2005;54:1421-7

70. Armuzzi A, Ahmad T, Ling KL *et al*. Genotype-phenotype analysis of the Crohn's disease susceptibility haplotype on chromosome 5q31. Gut. 2003;52:1133-9

71. Vermeire S, Pierik M, Hlavaty T *et al*. Association of organic cation transporter risk haplotype with perianal penetrating Crohn's disease but not with susceptibility to IBD. Gastroenterology. 2005;129:1845-53.

72. Noble CL, Nimmo ER, Drummond H *et al*. The contribution of OCTN1/2 variants within the IBD5 locus to disease susceptibility and severity in Crohn's disease. Gastroenterology. 2005;129:1854-64

73. Wakabayashi M, Ito T, Mitsushima M *et al*. Interaction of lp-dlg/KIAA0583, a membrane-associated guanylate kinase family protein, with vinexin and beta-catenin at sites of cell-cell contact. J Biol Chem. 2003;278:21709-14.

74. Friedrichs F, Brescianini S, Annese V *et al*. Evidence of transmission ratio distortion of DLG5 R30Q variant in general and implication of an association with Crohn disease in men. Hum Genet. 2006;119:305-1

75. Yamazaki K, Takazoe M, Tanaka T *et al*. Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. J Hum Genet 2004;49:664-8.

76. Buning C, Geerdts L, Fiedler T *et al*. DLG5 Variants in Inflammatory Bowel Disease Am J Gastroenterol. 2006 Feb 22; [Epub ahead of print]

77. Noble CL, Nimmo ER, Drummond H *et al*. DLG5 variants do not influence susceptibility to inflammatory bowel disease in the Scottish population. Gut. 2005 Oct;54:1416-20

78. Tremelling M, Waller S, Bredin F *et al*. Genetic Variants in TNF-alpha But Not DLG5 Are Associated with Inflammatory Bowel Disease in a Large United Kingdom Cohort. Inflamm Bowel Dis. 2006 ;12:178-184

79. Ho GT, Moodie FM, Satsangi J. Multidrug resistance 1 gene (P-glycoprotein 170): an important determinant in gastrointestinal disease? Gut 2003;52:759-66

80. Panwala CM, Jones JC, Viney JL *et al*. A novel model of inflammatory bowel disease: mice deficient for the multiple drug resistance gene, mdr1a, spontaneously develop colitis.J Immunol. 1998;161:5733-44

81. Ho GT, Nimmo ER, Tenesa A *et al*. Allelic variations of the multidrug resistance gene determine susceptibility and disease behavior in ulcerative colitis.Gastroenterology. 2005;128:288-96.

82. Potocnik U, Ferkolj I, Glavac D *et al*. Polymorphisms in multidrug resistance 1 (MDR1) gene are associated with refractory Crohn disease and ulcerative colitis. Genes Immun. 2004;5:530-9.

83. Brant SR, Panhuysen CI, Nicolae D *et al*. MDR1 Ala893 polymorphism is associated with inflammatory bowel disease. Am J Hum Genet. 2003;73:1282-92.

84. Croucher PJ, Mascheretti S, Foelsch UR *et al*. Lack of association between the C3435T MDR1 gene polymorphism and inflammatory bowel disease in two independent Northern European populations. Gastroenterology. 2003;125:1919-20

85. Glas J, Torok HP, Schiemann U *et al*. MDR1 gene polymorphism in ulcerative colitis. Gastroenterology. 2004;126:367.

86. Oostenbrug LE, Dijkstra G, Nolte IM *et al*. Absence of association between the Multidrug Resistance (MDR)1 gene and inflammatory bowel disease Scand J Gastroenterol. 2006;41:1174-82.

87. Cummings JR, Jewell DP. Clinical implications of inflammatory bowel disease genetics on phenotype. Inflamm Bowel Dis. 2005;11:56-61

88. Reumaux D, Sendid B, Poulain D Duthilleul P, Dewit O, Colombel JF. Serological markers in inflammatory bowel diseases. Best Pract Res Clin Gastroenterol. 2003;17:19-35.

89. Gasche C, Scholmerich J, Brynskov J *et al*. A simple classification of Crohn's disease report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. Inflamm Bowel Dis. 2000;6:8-15.

90. Oostenbrug LE, van Dullemen HM, te Meerman GJ *et al*. Clinical outcome of Crohn's disease according to the Vienna classification: disease location is a useful predictor of disease course. Eur J Gastroenterol Hepatol. 2006;18:255-61.

91. Silverberg MS, Satsangi J, Ahmad T *et al*. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol. 2005;19 Suppl A:5-36.

92. Sachar DB, Bodian CA, Goldstein ES *et al*. Task Force on Clinical Phenotyping of the IOIBD. Is perianal Crohn's disease associated with intestinal fistulization? Am J Gastroenterol. 2005;100:1547-9.

Chapter 3

Association of interleukin 1 receptor-associated kinase M *(IRAK-M)* and inflammatory bowel diseases

R.K. Weersma, L.E. Oostenbrug, I.M. Nolte,
G. van der Steege, E. Oosterom, H.M. van Dullemen,
J.H. Kleibeuker, G. Dijkstra

Scandinavian Journal of Gastroenterology. 2007;42:827-33

Abstract

Background: Inflammatory bowel diseases (IBD) have a complex genetic background. The Interleukin Receptor associated Kinase-M (IRAK-M) is a NF- κB -mediated, negative regulator of Toll-like receptor (TLR) signaling. A functional mutation in a negative regulator might induce impaired endotoxin tolerance and increased inflammatory responses. *IRAK-M* is situated on chromosome 12q14, a susceptibility locus for IBD, what makes it a good candidate gene. We analyzed a large cohort of IBD patients for the association of IBD with *IRAK-M*.

Methods: 542 patients with IBD (309 Crohn's disease (CD), 233 Ulcerative Colitis (UC)) and 305 controls were studied. Two single nucleotide polymorphisms (V147I and V270I) and six microsatellite markers were evaluated by association analysis and the Haplotype Sharing Statistic. Results were stratified for *CARD15* mutations R702W, G908R and 1007fsinsC.

Results: No significant differences in *IRAK-M* allele frequencies were observed between IBD, UC, CD or subgroups of CD or UC and controls. Five of 36 UC patients (13.9%) with an IBD associated *CARD15* mutation were carrier versus 2/167(1.2%) in non-carriers (OR13.1, 95%CI 1.0-164.5). No interaction was observed for CD.

Conclusion: No evidence was found for an association between IBD, CD, UC or subsets of CD and UC and *IRAK-M*. However, interaction was found between *IRAK-M* and *CARD15* in UC patients. In *CARD15* mutant patients the production of *IRAK-M* upon stimulation might be impaired. Further studies are needed to test if an impaired negative regulation of the TLR-signaling pathway might be partly responsible for the development of IBD.

Introduction

Crohn's disease (CD) and ulcerative colitis (UC) are the major forms of chronic inflammatory bowel diseases (IBD) and are characterized by chronic relapsing inflammation of the gastrointestinal tract. The combined prevalence of CD and UC is estimated at 100 – 200 / 100.000 in developed countries.[1 2]

Concordance rates in twins and siblings suggest that a genetic predisposition contributes to the pathogenesis of IBD.[3-5] Inflammatory bowel diseases do not follow a simple Mendelian mode of inheritance and the penetrance and expression of IBD is variable.[6] Both CD and UC are considered complex genetic disorders. Genetic linkage analyses, through genome wide screens, have identified a number of susceptibility loci and led to the discovery of the association of CD with the caspase activation recruitment domains gene on chromosome 16q.[7-9]

The main factor in the pathogenesis of IBD is a dysregulated immune response to intraluminal antigens. *CARD15* is an intracellular component of the innate immune system and is involved in the recognition of microbial products consisting of pathogen-associated molecular patterns (PAMPs) and activation of Nuclear Factor kappa B (NF- κB). Besides statistical evidence of a genetic association with IBD, *CARD15* has also been shown to have a functional role in activating NF- κB.[10] The precise mechanism how functional mutations in the NOD2/*CARD15* gene and subsequently an impaired activation of NF- κB leads to susceptibility of IBD is still only partially understood.

Toll-like receptors (TLRs) are, like *CARD15*, important components of the innate immune system. TLRs have extracellular leucine-rich-repeat domains that recognize PAMPs and activate adaptive immunity.[11] Chronic repeated stimulation of TLRs by PAMPs causes an "endotoxin tolerance".[12] This process involves of different mechanisms, including TLR4 downregulation and decreased NF- κB activation.[13-15] The interleukin-receptor associated kinase-M (IRAK-M, IRAK-3) is a NF- κB dependent, negative regulator of Toll-like receptor signaling pathway. Genetically engineered *IRAK-M* knockout mice have an impaired endotoxin tolerance and showed increased inflammatory responses to bacterial infection.[16 17] The gene coding for *IRAK-M* is on chromosome 12q14.2, and co-localizes with a susceptibility locus for IBD (IBD2), that was identified by a two-stage genome-wide search for inflammatory bowel diseases.[18]

The function of IRAK-M as a negative regulator of TLR signaling, inducing an impaired endotoxin tolerance and an inflammatory response, in combination with it's localization at the IBD2 locus makes *IRAK-M* a good candidate gene for susceptibility for inflammatory bowel disease. Therefore, we studied the association of *IRAK-M* with IBD and with specific UC and CD phenotypes and the interaction with *CARD15* in a cohort of patients from the northern part of the Netherlands.

Table 1. Marker data.

Locus	Marker Name	Position	Heterozygosity	Forward Primer (5'->3')	Reverse Primer (3'->5')
1	A	-5kb 5'	56.2	FAM-GAGTGTAACATCACAC-CCAGCTA	GCATTCCTCCCTGTCTAGGTTA
2	B	15 kb	77.1	GCCAAGGCTCTTAGATTGGAAAC	HEX-CAATGTGAAGCTACAGTTTTCTGG
3	V147I	25 kb	10.7	TTCTTAGCTAACTTATCTTC-CCCCTACTT	TTCCTTCTATGATATTTTGAAAGCTGAT-GGA
4	C	27 kb	89.6	GACTTTGCCCTTCTGCTTTTAT	HEX-AATGTTCTGTGGCTAAGGCTTC
5	V270I	40 kb	100.0	CCCACTCCCTTGGCACA	GCAGGTAGTGAATGGCTTTGGATAT
6	D	44 kb	10.2	GGAGGAGTGAACAATGGAAGAT	FAM-AAACAGCAGCATAGGGCTTTTA
7	D12S288	+ 27 kb 3'	52.3	GGATAATGCCCTGTACCCAGTA	HEX-GCAGACCAGACAGTAGCACATC
8	F	+ 40 kb 3'	38.3	FAM-CCAGTAACTCCCGTTGGG-TAAT	GCCATTCCAAATCTATCTGTCG

Methods

Patients and controls

Patients were recruited from the outpatient department of the University Medical Center Groningen, the Netherlands. Patients fulfilling diagnostic criteria for CD an UC were eligible for the study. These criteria are based on accepted clinical, endoscopic, radiological and histological findings.[19] For CD patients, phenotypic details were obtained according to the Vienna Classification.[20] For UC patients, phenotypes were described according to age of onset, extend of disease (proctitis, leftsided, or extended), necessity of colectomy and the occurrence of malignancy and extraintestinal manifestations. Duration of follow up was defined as the interval between diagnosis and the moment of case note review. The Medical Ethics Committees of the University Medical Center Groningen approved of the study. Written informed consent was obtained from all participating subjects. All DNA samples and data in this study were handled anonymously and individuals were aware that they would not be informed about individual test results. When DNA of parents was available the non-transmitted haplotypes of each parent were used as a control. When DNA of a child and a spouse was available, both haplotypes of the spouse were regarded as a control. In case of missing data, all available family members (i.e. parents, sibs or children of the patients) were used to reconstruct the missing data.

Genotyping and SNP selection

DNA was extracted from 20 ml EDTA-blood following standard procedures and was stored at -80 oC. Two known single nucleotide polymorphisms (SNP) and six microsatellite markers in and around the gene encoding for IRAK-M were genotyped. One SNP in exon 5, V147I (rs1152888) and a SNP in exon 8, V270I (rs11465972) were studied. One microsattelite marker was the known D12S288, located 27 kb downstream of IRAK-M. The other five microsatellite loci were found on the genomic sequence segment comprising the IRAK-M gene, with a margin of 40 kb downstream and upstream, and downloaded from the NCBI's public database. Marker details are given in Table 1. For CARD15 the three known mutations R702W, G908R, and 1007fsinsC were evaluated.

Primers to amplify the polymorphic loci were selected using on-line Primer3 software.[21] SNP genotyping was carried out by using TaqMan PCR primer/probe sets, designed through Applied Biosystems' Assay by Design service (http://myscience.appliedbiosystems.com/, Foster City, USA). SNP assay reactions were performed in 5 µl volumes and contained 25 ng DNA, 1x TaqMan Universal PCR Master Mix (Applied Biosystems), 100 nM of each primer and 900 nM of each probe. Cycling conditions on the ABI prism 7900 HT (Applied Biosystems) were 2 minutes 50°C, 10 minutes 95°C followed by 40 cycles of 15 seconds 92°C and 1 minute 60°C. End-point fluorescence was measured immediately after cycling. Alleles were assigned using SDS 2.0 software (Applied Biosystems).

For each microsatellite marker polymerase chain reaction (PCR), 0.5 units Taq DNA polymerase (Amersham Pharmacia Biotech, Uppsala, Sweden) was used to amplify the fragments. Reaction mixtures contained 0.2 mM dNTP (Roche Diagnostics, Mannheim, Germany), 2.5 mM MgCl2, 10 mM Tris-HCl (pH 9.0), 50 mM KCl (Amersham Pharmacia Biotech) and 0.25 µM

of each primer (with one primer 5' labeled with a fluorochrome 6-FAM or HEX (Sigma, Malden, the Netherlands). Cycling was done on a PTC-225 thermal cycler (MJ Research, Waltham, MA) and a PrimusHT (MWG Biotech, Ebersberg, Germany). A standard protocol was used for amplification. Post PCR multiplexing was performed by combining 2-10 µl (based on signal strength) of the PCR products. A 2.3 µl sample of the pooled fragments was mixed with 2.5 µl MilliQ and 0.2 µl ET-400R size standard (Amersham Pharmacia Biotech) and separated on a MegaBACE 1000 capillary sequencer (Amersham Pharmacia Biotech) according the manufacturers protocol. Results were analyzed using Genetic Profiler version 2.0 (Amersham Pharmacia Biotech). Scoring of the alleles was blinded for affection status and family structure.

Statistical methods

After genotyping the markers, Hardy-Weinberg equilibrium was tested among the unrelated unaffected individuals of the sample and inheritance errors were determined. If a marker showed deviation from Hardy-Weinberg equilibrium or the fraction of erroneous inheritance exceeded 5%, the marker was discarded from further analyses.

As our sample includes a large cohort of trios, haplotypes could be constructed. The non-transmitted haplotypes of the parents or the haplotypes of the spouse of the patient served as control haplotypes.

The allele frequencies of patients and controls were compared to test for association using the chi-square test or the Fisher's exact test when appropriate. Differences between the haplotypes of patients and the control haplotypes were also analyzed by Haplotype Sharing Statistics (HSS), which analyses length of haplotype similarity. The validity of this method has been demonstrated elsewhere by applying it extensively to simulated data as well as empirical data.[22-27] Haplotype analysis assumes that in genomic regions flanking disease mutations, haplotypes of patients will be more similar to each other and more extended than haplotypes of controls.

As a test of Linkage Disequilibrium (LD) in order to prove the presence of conserved haplotypes, the D' for multi-allelic markers was used.[28] The significance of the observed D' value was determined by the fraction of 1,000 randomizations (for which alleles were randomly redistributed over the haplotypes independently for all loci) that revealed a larger D' value than the observed one.

A multiple testing correction was performed for the number of complementary subgroups of patients using a Bonferroni correction. For allele and genotype associations, all patients were included. For the HSS, which require phase to be derived, only those with at least one participating family member were used.

Results

A total of 542 patients with IBD (309 with CD and 233 with UC) and 305 controls were included. Clinical characteristics of patients with CD and UC were available from 520 patients (table 2 and 3). For haplotype analysis 403 trios or duos were available. The percentage of

Table 2. Characteristics of patients with Crohn's disease.

Total number	309		
Sex (m/f)	108 / 201		
Age at diagnosis (years)			
Mean (range)	30.9 (6.7 – 73.9)		
Median	27.4		
Follow up (years)			
Mean (range)	11.9 (0.4 – 49.4)		
Median	9.0		
Disease localization			
Ileal	80	(25.9%)	
Colon	64	(20.7%)	
Ileocolon	165	(53.4%)	
Disease behavior			
Non-stricturing, non penetrating	100	(32.4%)	
Stricturing	75	(24.3%)	
Penetrating	134	(43.3%)	
Perianal	89	(28.8%)	
Extraintestinal manifestations	43	(13.9%)	
Family history of IBD	65		(21.0%)

Table 3. Characteristics of patients with ulcerative colitis.

Total number	233		
Sex (m/f)	122 / 111		
Age at diagnosis (years)			
Mean (range)	33.0 (20 – 46.0)		
Median	31.4		
Below40	166	(71.2%)	
Follow up (years)			
Mean (range)	13.7 (4.9 - 22.5)		
Median	11.9		
Disease localization ¶			
Proctitis	16	(6.8%)	
Leftsided	75	(32.2%)	
Pancolitis	120	(51.5%)	
Extra-intestinal manifestations	39	(16.7%)	
Family history of IBD	39		(16.7%)
History of surgical intervention	50	(21.4%)	

¶ Data unavailable for 22 (9,4%) patients

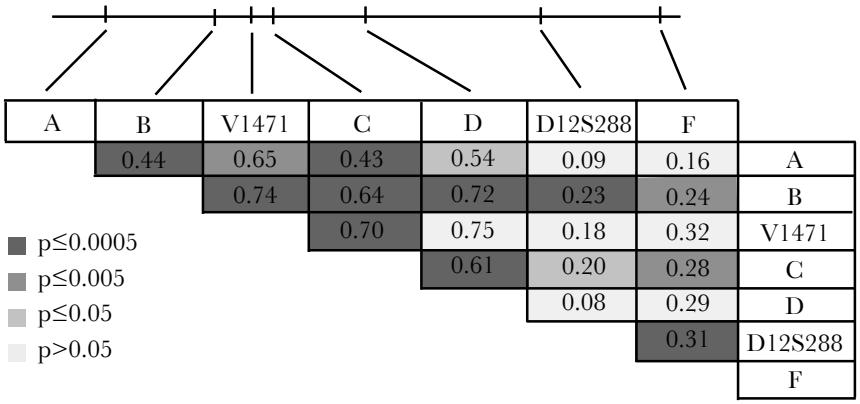


Figure 1. Linkage disequilibrium.
LD across *IRAK-M* and flanking genomic region. D' values for LD are shown for each pair of markers. The upper line depicts localization of markers across the genomic region. Most markers are in LD.

unknown alleles due to inheritance errors or PCR failure was 2.9%. Most markers were in LD (figure 1). SNP V270I in the encoding region for *IRAK-M* appeared not to be polymorphic in our sample; therefore this SNP was not informative for further analysis. All markers proved to be in Hardy-Weinberg equilibrium and to show a correct inheritance pattern in more than 97% of the trios.

Association analysis and Haplotype Sharing Statistic IRAK-M

No significant differences in the distribution of *IRAK-M* allele frequencies between the patients with IBD, CD or UC and controls were observed. Analysis of clinical subgroups of CD and UC patients did not reveal any significant differences. SNP V147I had a frequency for allele A of 5.6% for controls compared to 5.7% for IBD, 5.6% for CD and 5.9% for UC. A protective effect of allele 160 at marker F was observed for colonic localization and for stricturing behavior in CD. 11.3% of the controls, but only 1.4% of the patients with stricturing behavior and 1.7% of the patients with a colonic localization carried allele 160 homozygously. This effect did not reach statistical significance (p=0.11 for stricturing disease versus controls and p=0.21 for colonic localization versus controls).

Similar to the association analysis, the HSS revealed no difference in haplotype sharing between patients and controls for IBD. Analysis of the subgroups with UC and CD and the different phenotypes of patients with CD did not show any difference with the HSS either.

Interaction IRAK-M and CARD15

The prevalence of one of the three risk associated mutations in *CARD15* was 27.1% in IBD, 34,1% in CD, 17.4% in UC and 18.3% in controls. For CD patients no interaction between

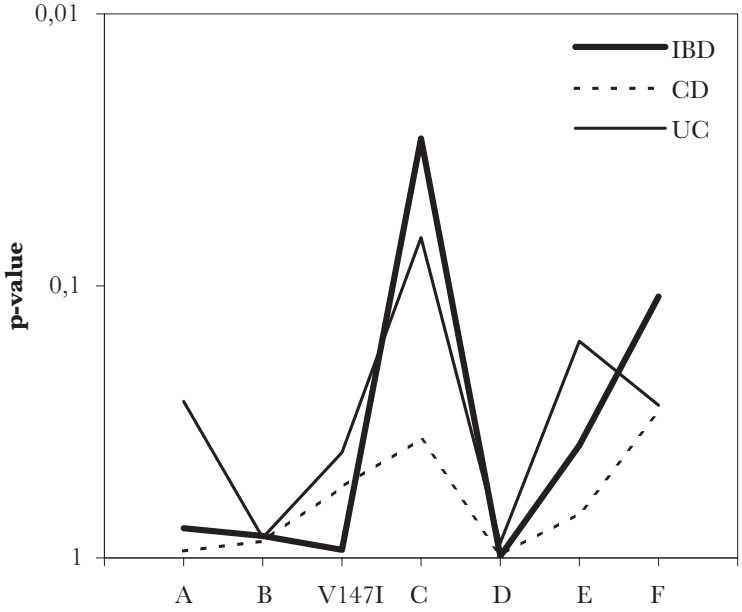


Figure 2. Allelic Association of *IRAK-M* and IBD in *CARD15* Carriers.
Association of *IRAK-M* and IBD, CD and UC, in carriers one of the three IBD associated mutations in *CARD15* (R702W, G908R, and 1007fsinsC), versus non-carriers.

CARD15 and *IRAK-M* was observed. For UC, interaction was observed between microsattelite marker C and *CARD15* status (Figure 2). Five of 36 UC patients (13.9%) with a *CARD15* mutation were carrier versus 2/167 (1.2%) not carrying an IBD associated mutation in *CARD15*. (OR 13.1, 95% confidence interval 1.0-164.5).

Discussion

We studied the association of IRAK-M with IBD and with specific CD phenotypes in a large cohort of patients from the northern part of the Netherlands. No evidence was found for an association between IBD, CD, UC or different subsets of CD and the IRAK-M gene region, indicating that mutations in IRAK-M are not involved in IBD susceptibility.
The pathogenesis of inflammatory bowel diseases still remains elusive. Mutations in the encoding gene for *CARD15* are associated with susceptibility for CD, but how these mutations are involved in the pathogenesis of IBD is poorly understood. *CARD15* is present in the cytoplasm of monocytes and epithelial cells and patients with *CARD15* mutations will have less NF- κB

activation upon stimulation with muramyl dipeptides (MDP). On the one hand, this impaired NF- κB activation could lead to a defective immune response against luminal bacteria.[29]
On the other hand, it is suggested in a recent study that *CARD15* functions as a negative regulator of TLR 2-mediated T helper type 1 responses.[30] *CARD15* (-/-) mice showed increased TLR 2-driven activation of NF- κB, particularly of the NF- κB subunit c-Re1, whereas intact NOD2 signalling inhibited TLR 2-driven activation of NF- κB. This is the first study that shows increased NF- κB activation and inflammatory responses in *CARD15* knockout mice, rather than less activation as shown in earlier models. This could explain why there is an increased susceptibility for CD in patients lacking functional *CARD15*.
IRAK-M is known as a negative regulator of the TLR signaling pathway.[16] IRAK-M is probably involved in a negative feedback loop, where TLR mediated activation of NF- κB will yield IRAK-M, which blocks further TLR activation, making the cell less responsive for TLR signaling.[31] The encoding gene is situated on chromosome 12q14, a region which was previously identified as a susceptibility locus (IBD2) for IBD.[18] We hypothesized that a mutation of such a negative regulator could explain the increased NF- κB activation as seen in *CARD15* wildtype and even in *CARD15* mutant patients. Gene products with these features are good candidate genes for susceptibility for IBD.
We studied a large cohort of IBD patients and family-based controls from the northern part of the Netherlands. Two known SNPs and six microsatellite markers spanning the entire *IRAK-M* encoding region were used. We failed to show an association of *IRAK-M* with IBD or with the clinical subsets CD or UC, both at a single locus level as well at a haplotype level. No association with different phenotypes in CD or UC was found either. Power analysis showed that an association with an odds ratio of >1.75 should have been detected in our study population. Nevertheless, due to small size of the different subsets a smaller effect could have been missed.
No interaction of *IRAK-M* and *CARD15* was observed in CD patients or different subsets of CD patients. However, carriership of one or more of the three IBD associated mutations in *CARD15* in combination with microsattelite marker C showed an increased risk for UC compared to patients not carrying a *CARD15* mutation. Since *CARD15* has not been shown to be involved in UC susceptibility before, it remains doubtful whether this is a true association or a false positive result.
Marker F revealed an interesting protective association for stricturing or colonic disease in Crohn's disease but this was not statistically significant. Marker F is situated 40kb downstream of the coding region for *IRAK-M* in an area coding for LOC390340, a protein similar to phosducin-like 3 with no known function to date. Because of the weak association and the lack of a functional explanation, no further investigations were done in this region.
In conclusion, in this Dutch cohort we could not find an association between IBD and *IRAK-M* at the SNP and haplotype level. We did find an association of *IRAK-M* and Ulcerative Colitis in carriers of a *CARD15* mutation compared to non-carriers, however this was based on small numbers. The lack of a genetic association does not rule out that in *CARD15* mutant patients the production of IRAK-M upon MDP stimulation is impaired, resulting in increased TLR signaling. Further studies on the regulation of TLR signaling in *CARD15* mutant patients and on the role of IRAK-M in IBD are needed to test this hypothesis. A defect in the negative regulation of TLR signaling, inducing an impaired endotoxin tolerance and an increased inflammatory response might be partly causative in IBD.

References

1. Binder V, Both H, Hansen PK, Hendriksen C, Kreiner S, Torp-Pedersen K. Incidence and prevalence of ulcerative colitis and Crohn's disease in the County of Copenhagen, 1962 to 1978. Gastroenterology 1982;83:563-68.

2. Calkins BM, Mendelhoff AI. The epidemiology of idiopathic inflammatory bowel diseases. Kirsner JB and Shorter RG, eds. Inflammatory Bowel Diseases. 31-68. 1995. Baltimore, Williams & Wilkins.

3. Orholm M, Munkholm P, Langholz E Nielsen OH, Sorensen TI, Binder V.Familial occurrence of inflammatory bowel disease. N Engl J Med 1991;324:84-88.

4. Halfvarson J, Bodin L, Tysk C, Lindberg E, Jarnerot G. Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. Gastroenterology 2003;124:1767-73.

5. Laharie D, Debeugny S, Peeters M *et al.* Inflammatory bowel disease in spouses and their offspring. Gastroenterology 2001;120:816-19.

6. Orholm M, Iselius L, Sorensen TI Munkholm P, Langholz E, Binder V. Investigation of inheritance of chronic inflammatory bowel diseases by complex segregation analysis. BMJ 1993;306:20-24.

7. Hugot JP, Laurent-Puig P, Gower-Rousseau C *et al.* Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature 1996;379:821-23.

8. Hugot JP, Chamaillard M, Zouali H *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 2001;411:599-603.

9. Ogura Y, Bonen DK, Inohara N *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 2001;411:603-6.

10. Ogura Y, Inohara N, Benito A Chen FF, Yamaoka S, Nunez G. Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB. The Journal of biological chemistry 2001;276:4812-18.

11. Medzhitov R, Janeway CA Jr. Innate immunity: the virtues of a nonclonal system of recognition. Cell 1997;91:295-98.

12. Ziegler-Heitbrock HW. Molecular mechanism in tolerance to lipopolysaccharide. J Inflam 1995;45:13-26.

13. Nomura F, Akashi S, Sakao Y *et al.* Cutting edge: endotoxin tolerance in mouse peritoneal macrophages correlates with down-regulation of surface toll-like receptor 4 expression. J Immunol 2000;164:3476-79.

14. Kastenbauer S, Ziegler-Heitbrock HW. NF-kappaB1 (p50) is upregulated in lipopolysaccharide tolerance and can block tumor necrosis factor gene expression. Infect Immun 1999;67:1553-59.

15. Goldring CE, Reveneau S, Pinard D Jeannin JF. Hyporesponsiveness to lipopolysaccharide alters the composition of NF-kappaB binding to the regulatory regions of inducible nitric oxide synthase gene. Eur J Immun 1998;28:2960-2970.

16. Kobayashi K, Hernandez LD, Galan JE Janeway CA Jr, Medzhitov R, Flavell RA. IRAK-M is a negative regulator of Toll-like receptor signaling. Cell 2002;110:191-202.

17. Nakayama K, Okugawa S, Yanagimoto S *et al.* Involvement of IRAK-M in peptidoglycan-induced tolerance in macrophages. J Biol Chem 2004;279:6629-34.

18. Satsangi J, Parkes M, Louis E *et al.* Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nat Gen 1996;14:199-202.

19. Podolsky DK. Inflammatory bowel disease. N Engl J Med 2002;347:417-29.

20. Gasche C, Scholmerich J, Brynskov J *et al.* A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. Inflamm Bowel Dis 2000;6:8-15.

21. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365-86.

22. Beckmann L, Fischer C, Deck KG *et al.* Exploring haplotype sharing methods in general and isolated populations to detect gene(s) of a complex genetic trait. Genet Epidemiol 2001;21 Suppl 1:S554-S559.

23. Boon M, Nolte IM, Bruinenberg M *et al.* Mapping of a susceptibility gene for multiple sclerosis to the 51 kb interval between G511525 and D6S1666 using a new method of haplotype sharing analysis. Neurogenetics 2001;3:221-30.

24. Levinson DF, Kirby A, Slepner S Nolte I, Spijker GT, te Meerman G. Simulation studies of detection of a complex disease in a partially isolated population. Am J Med Genet 2001;105:65-70.

25. Nolte IM, Spijker GT, Boon M *et al.* The Haplotype Sharing Statistic: fine-mapping of disease gene loci by comparing patients and controls for the length of haplotype sharing. Thesis 2002.

26. de Jong MM, Nolte IM, de Vries EG *et al.* The HLA class III subregion is responsible for an increased breast cancer risk. Hum Mol Genet 2003.

27. Oostenbrug LE, Drenth JP, de Jong DJ *et al.* Association Between Toll-like Receptor 4 and Inflammatory Bowel Disease. Inflamm Bowel Dis 2005;11:567-75.

28. Lewontin R. The interaction of selection and linkage. I. General considerations; heterotic models. Genetics 1964;49:49-67.

29. Rogler G. Update in inflammatory bowel disease pathogenesis. Curr Opin Gastroenterol 2004;20:311-17.

30. Watanabe T, Kitani A, Murray PJ Strober W. NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses. Nat Immunol 2004;5:800-808.

31. Escoll P, del Fresno C, Garcia L, Valles G, Lendinez MJ, Arnalich F, Lopez-Collazo E. Rapid up-regulation of IRAK-M expression following a second endotoxin challenge in human monocytes and in monocytes isolated from septic patients. Biochem Biophys Res Commun 2003;311:465-72.

Chapter 4

The Runt-Related Transcription Factor 3 *(RUNX3)* is associated with ulcerative colitis and shows epistasis with Solute Carrier Family 22, members 4 and 5 *(SLC22A4/5)*

R.K. Weersma, L. Zhou, I.M. Nolte, G. van der Steege, H.M. van Dullemen, E. Oosterom, L. Bok, K.N. Faber, M.P. Peppelenbosch, J.H. Kleibeuker, G. Dijkstra

Submitted

Abstract

Background: Inflammatory bowel diseases (IBD) –comprising Crohn's disease(CD) and ulcerative colitis(UC)– are intestinal inflammatory disorders with a complex genetic background. Mice deficient for the runt-domain-transcription factor 3 (Runx3) develop spontaneous colitis. Human *RUNX3* resides in a susceptibility locus for IBD. We studied the association of *RUNX3* in a cohort of IBD patients and analyzed the interaction with *SLC22A4/5*. RUNX3 and OCTN1 mRNA expression was assessed in inflamed and non-inflamed mucosa from patients and controls.

Methods: 543 IBD patients (309 CD/234 UC) and 296 controls were included. Four Single Nucleotide Polymorphisms (SNPs) and four microsatellite markers were studied for *RUNX3*. Five SNPs (including SNP-207G→C and SNP1672C→T) were analyzed for *SLC22A4/5*. RUNX3 and OCTN expression in mucosal tissue from 30 patients (14UC/16 CD) and 6 controls was determined by Q-PCR.

Results: We demonstrate significant association between *RUNX3*-SNP rs2236851 and UC (OR 1.61;CI 1.11-2.32,p=0.020). Carriership is associated with pancolitis (OR 1.86;CI 1.08-3.21). *SLC22A4/5*-SNPs rs272893 and rs273900 are associated with CD (OR 2.16;CI 1.21-3.59 and OR 2.40;CI 1.43-4.05). We find epistasis for carriership of a risk-associated allele in *RUNX3* and *SLC22A4/5* for UC patients versus CD patients(OR 3.83;CI 1.26-11.67). RUNX3 mRNA expression is increased ($p<0.05$) and OCTN1 expression is decreased ($p<0.05$) in inflamed colonic mucosa of IBD patients compared to their non-inflamed mucosa and controls.

Conclusions: We provide the first evidence that *RUNX3* is associated with UC and shows epistasis with *SLC22A4/5*. Additionally, we show a correlation between increased RUNX3 and decreased OCTN1 expression in inflamed mucosa of IBD patients.

Introduction

Chronic inflammatory bowel diseases (IBD) comprising Crohn's disease (CD) and ulcerative colitis (UC) are characterized by chronic relapsing inflammation of the gastrointestinal tract.[1] The combined prevalence of CD and UC is estimated at 100 – 200 / 100,000 in developed countries. Concordance rates in twins and siblings suggest that a genetic predisposition contributes to the pathogenesis of IBD. The *CARD15* gene on chromosome 16 was identified as the first susceptibility gene for IBD by performing genome-wide linkage studies. Three mutations (R702W, G908R, and 1007fsinsC) are independently associated with CD in Caucasian patients. Since IBD are complex genetic diseases with many different phenotypes, multiple genes are likely involved in different subsets of patients.

Functional mutations in Solute Carrier Family 22 members 4 and 5 (*SLC22A4* and *SLC22A5*) on chromosome 5q31 (IBD5) encoding for the Organic Cation Transporters 1 and (OCTN1/2) have been reported to be associated with CD. The risk haplotype consisting of the mutations 1672T→C in *SLC22A4* and -207G→C in *SLC22A5* had a higher frequency among CD patients (54%) than among controls (42%).[7 8] An epistatic effect of the risk haplotype with the *CARD15* gene was also reported. However, this has not been confirmed in a consecutive study.[9]

Besides genome-wide linkage studies, IBD candidate genes can also be selected from gene knock-out studies in mice. Disruption of single genes in mice may lead to the development of spontaneous colitis or an increased sensitivity to chemically-induced colitis, with similar phenotypes as observed for IBD. Mice that lack the Runt-related transcription factor 3 (Runx3) develop such a phenotype.[10] RUNX3 forms the runt domain family of transcription factors with RUNX1 and RUNX2, which share the same highly conserved runt domain. RUNX3 is involved in T-cell differentiation and functions as a tumour suppressor gene in gastric cancer.[11 12] It is highly expressed in dendritic cells (DCs) where it is involved in the Transforming Growth Factor-ß (TGF- ß) signaling pathway. The RUNX proteins appear to have a potential role in autoimmune diseases like rheumatoid arthritis, systemic lupus erythematosis and psoriasis although very little is known how these transcription factors are involved in these disorders.[13] The gene encoding for RUNX3 resides on the chromosomal region 1p36, which has been found to be a susceptibility locus for IBD.[14 15] Moreover, RUNX3 and TGF- ß are downregulated in peripheral blood cells of CD patients, which might suggest involvement of this pathway in the human pathogenesis of IBD[16] In view of these considerations *RUNX3* seems a good candidate gene for susceptibility for IBD. More specifically, regarding the development of a colitis similar to UC in knockout mice, it might be associated with ulcerative colitis.

Interestingly, a polymorphism in the RUNX binding site in *SLC22A4* was linked to rheumatoid arthritis.[17] Since this binding site is the same for all three members of the RUNX family, interaction between *SLC22A4/5* and *RUNX3* might be involved in IBD susceptibility.

We studied 543 IBD patients, in most cases participating with family members to form trios, for association of *RUNX3* and *SLC22A4/5* with IBD. Genetic epistasis between the two genes and separately with *CARD15* was analyzed. In addition, *in vitro* data on RUNX3 and OCTN mRNA expression in normal and inflamed colonic mucosa from patients with IBD are presented.

Table 1. Clinical characteristics of patients with Crohn's disease included in the genotyping cohort.

Total number	309	
Sex (m/f)	108 / 201	
Age at diagnosis (years)		
Mean (range)	30.9	(6.7 – 73.9)
Median	27.4	
Below 40	227	(73.5%)
Follow up (years)		
Mean (range)	11.9	(0.4 – 49.4)
Median	9.0	
Disease localization		
Ileal	80	(25.9%)
Colon	64	(20.7%)
Ileocolon	165	(53.4%)
Disease behavior		
Non-stricturing, non penetrating	100	(32.4%)
Stricturing	75	(24.3%)
Penetrating	134	(43.3%)
Upper GI tract	24	(7.8%)
Peri-anal	89	(28.8%)
Extra-intestinal manifestations	43	(13.9%)
Family history of IBD	65	(21.0%)
History of surgical intervention	186	(60.2%)

Methods

Patients and controls

For the genetic association study, a cohort of 543 IBD patients (309 CD and 234 UC) from the University Medical Center Groningen was used.[18] Diagnosis of IBD was established by accepted clinical, endoscopic and histopathological criteria.[1] For 183 (108 CD and 75 UC) patients DNA of both parents was also available. Patients with CD were phenotyped using the Vienna classification.[19] Patients with UC were phenotyped according to disease localization (proctitis, leftsided or pancolitis), extraintestinal manifestations, the development of malignancy or the need for colectomy. Clinical characteristics of patients with CD and with UC are presented in tables 1 and 2 respectively.

Mucosal tissue samples were available for 30 patients with IBD (14 UC and 16 CD) and for 6 controls referred for endoscopy for analysis of abdominal pain, without endoscopic and microscopic abnormalities at colonoscopy. [20] Characteristics of patients and controls are presented in table 3. Intestinal mucosal biopsy specimens were obtained, during endoscopy from inflamed and non-inflamed areas in the terminal ileum and throughout the colon of

Table 2. Clinical characteristics of patients with ulcerative colitis included in the_genotyping cohort.

Total number	234	
Sex (m/f)	122 / 112	
Age at diagnosis (years)		
Mean (range)	33.0	(20 – 46.0)
Median	31.5	
Below 40	167	(71.4%)
Follow up (years)		
Mean (range)	13.7	(4.9 - 22.5)
Median	11.9	
Disease localization †		
Proctitis	16	(6.8%)
Leftsided	76	(32.5%)
Pancolitis	120	(51.3%)
Extra-intestinal manifestations	39	(16.7%)
Family history of IBD	39	(16.7%)
History of surgical intervention	50	(21.4%)

† Data unavailable for 22 (9,4%) patients.

Table 3. Clinical characteristics of patients included in the mucosal biopsies cohort.

	Male/Female	**Mean Age (years)**	**Medication (without / with)**	
CD (n=16)	3 / 13	40	7 / 9	Azathioprine (4), Steroids (5), Mesalazine (2).
UC (n=14)	6 / 8	48	7 / 7	Azathioprine (3), Steroids (3) Mesalazine (7).
Controls (n=6)	5 / 1	41	6 / 0	None

Table 4. Marker data for genetic analysis *RUNX3* and *SLC22A4/5*.

Marker Name	MS/SNP*	Position	Type	
RUNX3				
RX3-01	MS	-75 kb	$(CA)_n$	FP†:GAGCAAGATCCTGCCACTACA
				RP: ATATGCCCTGTTGTCCCTCT
RX3-02	MS	-12 kb	$(CA)_n$	FP: GCCAAGGGCATCACTAGGTA
				RP: CAGGGGTCACGAATATCCAG
rs2236848	SNP		G→A	
rs2282719	SNP		T→G	
rs2236851	SNP		G→A	
RX3-03	MS	Intra 20kb	$(TTCA)_n$	FP: ACCAAGCTGAAAATGCCTTG
				RP: TCTGGTCCATCAAAGCGAGT
rs742230	SNP		C→T	
RX3-04	MS	+30 kb	$(CATC)_n$	FP: AATCGTCAGCCCATTCATTC
				RP: CATTGCTGCCCCTCACTACT
SLC22A4/5				
rs3792876	SNP		C→T	
rs2728932	SNP		C→T	
rs1050152	SNP		1672 G→C	
rs273900	SNP		C→T	
rs2631367	SNP		-207 G→C	

* SNP Single nucleotide polymorphism, MS Microsatellite
† FP Forward Primer; RP Reverse Primer

patients with active CD or UC. In controls, biopsies were obtained from the terminal ileum, ascending colon, transverse colon, and rectum. The study was approved by the institutional ethics review committee and all patients gave informed consent.

Genotyping

DNA was extracted from 20 ml EDTA-blood following standard procedures and was stored at -80 ºC. Primers to amplify the polymorphic loci were selected using Primer3 software.[21] SNP genotyping and microsatellite marker polymerase chain reaction (PCR) were performed as described previously.[18]

No informative Single Nucleotide Polymorphisms (SNPs) were found in the exonic regions of *RUNX3*. Therefore, four SNPs in intronic regions were determined for *RUNX3* (rs2236848, rs2282719, rs2236851 and rs742230). In addition, four microsatellite loci were found on the genomic sequence segment comprising the *RUNX3* gene in the NCBI's public database. These microsattelite markers are named RX3-01, RX3-02, RX3-03 and RX3-04 in the present study. For *SLC22A4/5* the previously found associated polymorphisms rs2631367 (-207 G→C), rs1050152 (1672 C→G), the polymorphism disrupting the RUNX1 binding site rs3792876 as well as three additional SNPs (rs272893 rs273900 and rs274551) were analyzed.[16 19] Details of

all markers are given in table 4.
SNP rs274551 in *SLC22A4* appeared not to be in Hardy-Weinberg equilibrium and was excluded from further analysis. The percentage of unknown alleles due to PCR failure or inheritance errors was 1.8% for *RUNX3* and 1.6% for *SLC22A4/5*. Data for the three known mutations R702W, G908R, and 1007fsinsC in *CARD15* were available from a previous study.[22]

RNA Isolation and Quantitative PCR

The collection of mucosal biopsies, isolation of total RNA and conditions for quantitative PCR to determine mRNA expression were described before.[20] RUNX3 and OCTN1 expression was determined using the assay-on-demand kit on the ABI PRISM 7700 (Applied Biosystems, Foster City, California, USA). Each sample was analyzed in duplicate. 18S levels were used as endogenous control.

Statistical methods

After genotyping, Hardy-Weinberg equilibrium was tested among the unrelated unaffected individuals of the sample and inheritance errors were determined. If a marker showed deviation from Hardy-Weinberg equilibrium or the fraction of erroneous inheritance exceeded 5%, the marker was discarded from further analyses. As our sample includes a large number of patients and family members, haplotypes could be constructed. For haplotype analysis 403 trios or duos were available. The non-transmitted haplotypes of the parents or the haplotypes of the spouse of the patient served as control genotypes.
For the microsatellite markers, allele frequencies in patients and controls were compared to test for association using the chi-square test or the Fisher's exact test when appropriate. For SNPs, the genotype frequencies were compared. Frequencies of haplotypes of SNPs were estimated by the expectation-maximization (EM)-algorithm and differences between patients and controls were analyzed by means of the likelihood ratio test. In trios consisting of the patients and both parents (n=183), association was also examined by the Transmission / Distortion test (TDT).
Odds ratios (ORs) and 95% confidence intervals (CI) were estimated using binary logistic regression (SPSS 12). Binary logistic regression was also used to determine gene-gene interaction effects with inclusion of the main effects of each gene in the model. Differences between the haplotypes of patients and the control haplotypes were analyzed by Haplotype Sharing Statistics (HSS), which analyses length of haplotype similarity. The validity of this method has been demonstrated elsewhere. [18 23]
As a test of Linkage Disequilibrium (LD) in order to prove the presence of conserved haplotypes, the D' for multi-allelic markers was used. The significance of the observed D' value was determined by the fraction of 10,000 randomizations (for which alleles were randomly redistributed over the haplotypes independently for all loci) that revealed a larger D' value than the observed one. A multiple testing correction was performed for the number of subgroups of patients using a Bonferroni correction. This applies not only to the p-values but also to the confidence intervals of the OR. For allele, genotype and haplotype associations, all patients were included. For the HSS, which require phase to be derived, only those with at least one participating family member were used.

Results

Association of RUNX3 and IBD

SNP rs2236851 in *RUNX3* was significantly associated with UC (p=0.020, OR 1.61 CI 1.11-2.32). The T-allele was present in 40.2% of the UC patients compared to 33.8% of CD patients and in 29.5 % of the controls. This association was confirmed by the TDT (allele T was 34 times transmitted and 14 times not transmitted, p=0.004) and HSS (p=0.012). (Fig 1.) Carriership of rs2236851 in UC was associated with pancolitis (OR 1.86 CI 1.08-3.21) and a tendency to association with an early age of onset (OR 1.59 (CI 0.98-2.57)).

For microsatellite marker RX-02, an association was found for colonic localization of CD (p=0.022). For the microsattelite marker RX3-01 an association with IBD was found (p=0.045). This association seemed to be stronger in CD (p=0.051) than in UC (p=0.17). With respect to subgroups of CD, RX3-01 was associated with an early age of onset (p=0.039) and ileocolonic localization (p=0.026).

All markers for *RUNX3*, except for all pairs with intra-genic marker RX3-03 and the pairs RX3-02 with rs2236851 and rs2236848 with RX3-04, were in LD with each other. For all other combinations, the p-value of the observed values of D' obtained by randomization were smaller than 0.05 and for the SNP combinations even smaller than 0.003, implying strong LD within the gene.

Association of SLC22A4/5 and IBD

Results of the association analysis for *SLC22A4/5* are shown in figure 2. SNPs rs272893 and rs273900 were significantly associated with CD (p=0.008 and 0.004 respectively) but not with UC. For homozygotes the ORs for CD of were 2.16 (1.21-3.59) for rs272893 and 2.40 (1.43-4.05) for rs273900 compared to controls. Heterozygotes were not at increased risk. Both SNPs were associated with an age of onset >40 yrs (p=0.0016/0.0007), ileocolonic localization (p=0.006/0.002), upper GI tract localization (p=0.007/0.003), non-stricturing non-penetrating behavior (p=0.020/0.011), and extra-intestinal manifestations (p=0.001 and <0.001). These SNPs are almost in complete LD with each other (r2=0.97). The TDT revealed no significant distortion to UC patients or to CD patients in general for either SNP (Figure 2). Haplotype association analysis supported the results found by single locus association, but did not show higher levels of significance.

We did not find any association for SNP rs2631367 (-207 G→C) and SNP rs1050152 (L503F) with UC, CD or subgroups of CD. The previously described risk haplotype TC was observed in our population in 41.1% of CD patients, 45.9% of UC patients and 44.4% of controls.

All SNPs in *SLC22A4/5* were in strong LD with each other. For all combinations of SNPs, the p-value of the observed values of D' obtained by randomization were smaller than 0.001.

Interaction of SLC22A4/5, RUNX3 and CARD15.

Among patients with homozygosity for one or both associated SNPs rs272893 or rs273900 in *SLC22A4*, 14 out of 25 UC patients (56%) were carriers of rs2236851 in *RUNX3* compared to 14 out of 51 (27.5%) of CD patients and 8 out of 26 (30.8%) of controls. Binary logistic regression analysis revealed an OR of 3.83 for UC patients versus CD patients to carry the combination

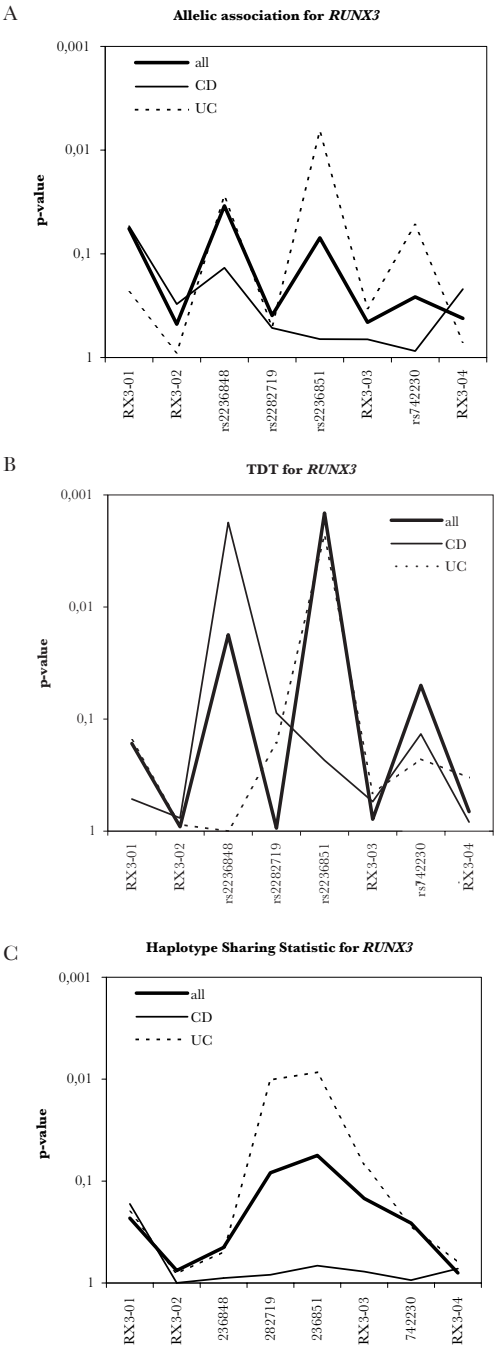


Figure 1.
Genetic association of *RUNX3* and IBD, CD and UC of 543 patients (309 CD, 234 UC) compared to healthy controls. SNP rs2236851 is associated with UC (OR 1.61 CI 1.11-2.32) but not with CD (1A). This association was confirmed by TDT (1B) and haplotype sharing statistics (1C).

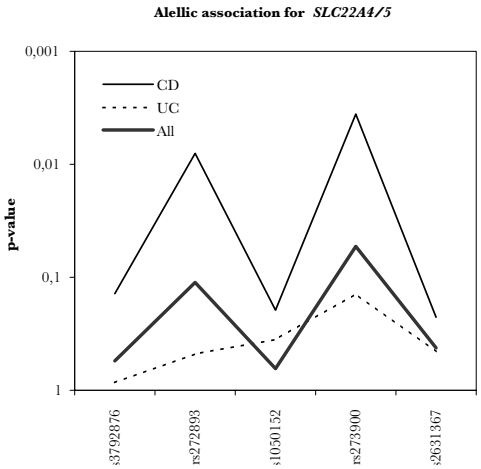
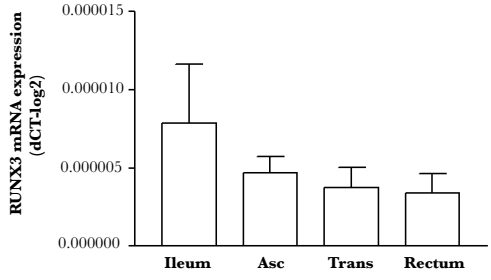


Figure 2.
Genetic association of Slc22a4/5 and IBD, CD and UC of 543 patients (309 CD, 234 UC) compared to healthy controls. SNPs rs272893 and rs273900 were associated with CD (OR 2.16 (1.21-3.59) and 2.40 (1.43-4.05)) on allelic association analysis.SNPs rs1050152 (1672C→T) and rs26313667 (-207G →C) were not associated with IBD, CD or UC.

A



B

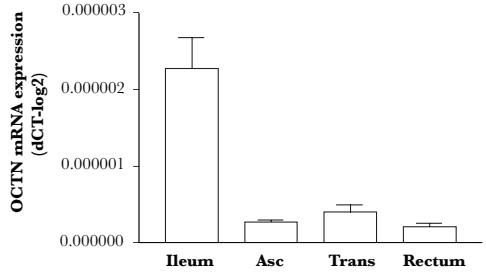


Figure 3.
Expression of RUNX3 (A) and OCTN1 (B) mRNA quantified by real time RT-PCR. RUNX3 expression is evenly distributed throughout the colon and the ileum. OCTN1 expression is increased in the ileum compared to the colon ($p<0.00001$). Total RNA was isolated from biopsies from colonic and ileal tissue from 6 controls. Relative expression levels (ΔCt) were normalized to 18S.Data are expressed as means ± SD.

Abbreviations: Asc -Ascending colon; Trans - Transverse colon

Table 5. Output of the binary logistic regression analysis for interaction of *RUNX3* and *SLC22A4/5* in Ulcerative Colitis (UC) versus Crohn's Disease (CD), showing an increased risk for UC in patients who have homozygosity for *SLC22A4/5* Allele A at rs272893 and/or *SLC22A4/5* allele T at rs273900 together with carriership of *RUNX3* allele T at rs2236851. Significance is corrected for multiple testing.

	p-value	OR	95.0% C.I.	
			lower	upper
SLC22A4/5 *	0.0082	0.36	0.17	0. 77
RUNX3 ¶†	0.73	1.07	0.72	1.61
SLC22A4/5 and RUNX3 *¶†	0.018	3.83	1. 26	11.67

* SLC22A4/5 defined as homozygosity for the risk allele of rs272893 and/or rs273900
¶ RUNX3 defined as carrying the risk allele of rs2236851
† Corrected for multiple testing

of homozygosity for allele A at rs272893 and/or allele T at rs273900 (i.e. risk factor for CD) and carriership of allele T at rs2236851 (i.e. risk factor for UC) (p=0.018, CI 1.26-11.67) (Table 5). No difference was observed between IBD patients and controls (OR=1.05 (CI 0.38-2.86)).
No significant interaction was found for the associated SNP rs2236851 in *RUNX3* and SNP rs3792876 which disrupts the RUNX binding site in *SLC22A4*. No evidence for epistasis of *CARD15* with *SLC22A4/5* or for *CARD15* with *RUNX3* was found in IBD, UC or CD or in subsets of UC or CD patients.

RUNX3 and OCTN1 mRNA expression in inflamed and non-inflamed intestinal mucosa

RUNX3 and OCTN1 mRNA expression in the colon and ileum of healthy controls is presented in figure 3a and 3b respectively. RUNX3 mRNA expression is evenly distributed throughout the colon and the ileum. OCTN1 expression is higher in the ileum compared to the colon (p<0.00001), while the expression level is constant throughout the colon.
Expression levels of RUNX3 mRNA in inflamed colonic mucosa of CD and UC patients were increased compared to controls (Figure 4). This difference was not observed in inflamed ileal mucosa in CD patients compared to controls. RUNX3 expression levels are increased in inflamed colonic mucosa compared to non-inflamed colonic mucosa in UC patients (p=0.01). OCTN1 expression is decreased in inflamed colonic mucosa compared to non-inflamed colonic mucosa (p=0.08) (figure 4a and 4b). In CD, RUNX3 is upregulated in inflamed intestinal mucosa (ileal and colonic localization combined) compared to non-inflamed mucosa (p=0.04) and OCTN1 down regulated (p=0.008).

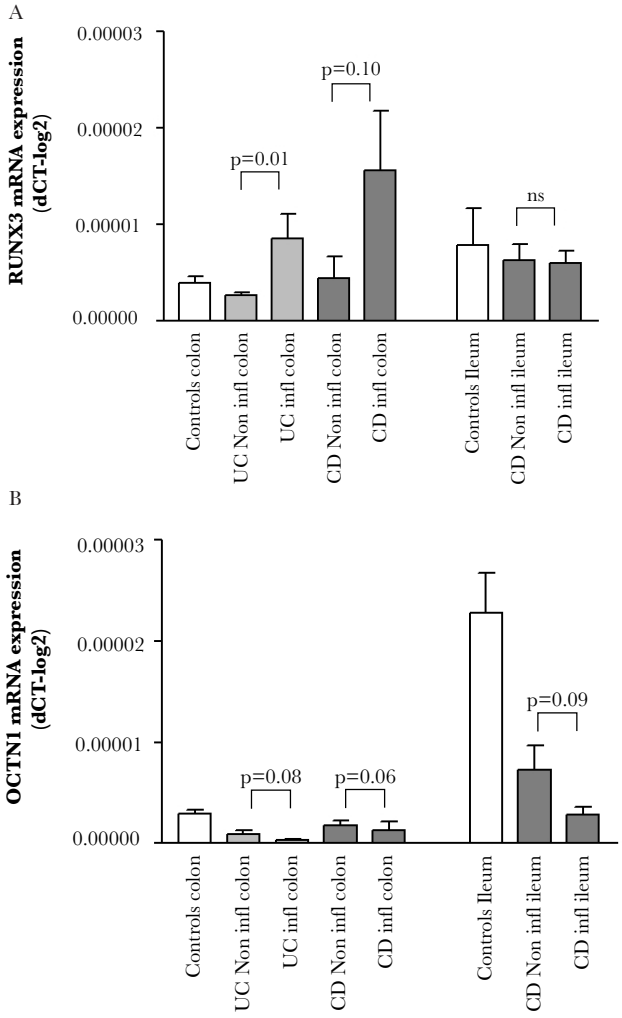


Figure 4.

Expression of RUNX3 (fig 4a) and OCTN1 (fig 4b) mRNA quantified by RT-PCR. RUNX3 mRNA is increased in inflamed colonic mucosa compared to non-inflamed mucosa in UC (p=0.01) and in CD (not significant). RUNX3 mRNA is not increased in inflamed ileal mucosa compared to non-inflamed ileal mucosa in CD. OCTN1 mRNA is decreased (but not statistically significant) in inflamed colonic mucosa compared to inflamed colonic mucosa in UC and CD and in inflamed ileal mucosa compared to non-inflamed ileal mucosa in CD. Total RNA was isolated from biopsies from uninflamed and inflamed colonic or ileal tissue from 30 IBD patients (14 UC and 16 CD) and 6 controls. Relative expression levels (ΔCt) were normalized to 18S. Data are expressed as means ± SD.

Abbreviations: UC - Ulcerative colitis; CD - Crohn's Disease; Non infl - Non inflamed; Infl - inflamed

Discussion

In this study, we applied a candidate gene approach to analyze the genetic association between the genes *RUNX3* and *SLC22A4/5* with IBD, UC, CD or subsets of CD and UC.
This is the first report showing an association of genetic variation in *RUNX3* with ulcerative colitis. Carriership of the T-allele of SNP rs2236851 was associated with UC with an OR of 1.61, which was supported by HSS and TDT. The association was most significant for extensive UC.
As a TGF-ß1 target, the RUNX3 protein is involved in TGF-ß1 mediated inhibition of T-cell proliferation and differentiation, as well as down regulation of macrophage activation and dendritic cell maturation. The important role of adequate TGF-ß1 signaling is apparent in TFG-ß1 knockout mice, which die soon after birth. Transgenic mice expressing dominant negative TGF-ßRII chains which are unresponsive to TGF-ß1 signaling develop severe colitis and pulmonary inflammation.[24 25] Furthermore Runx3 knockout mice developed indolent non ulcerative colitis characterized by epithelial hyperplasia, leukocyte infiltration, formation of B cell clusters and increased production of IgA.
SNP rs2236951 is located in an intron of the RUNX3 gene and at present it is unclear how this may be related to UC susceptibility. This is, however, not a unique finding, since an intronic SNP in RUNX1 has previously been shown to be associated with rheumatoid arthritis.[17] Furthermore, several studies have reported that important transcriptional regulatory elements are located in intronic regions.[26] Involved mechanisms might be hypermethylation of the promoter region of *RUNX3* or mislocalization of the protein into the cytoplasm, as observed in gastric cancer, which can both result in decreased activity of RUNX3.[28] Whether this particular SNP has a causative role in UC susceptibility or it is in LD with other functional mutations in the *RUNX3* gene needs further investigation.
Endoscopical and histological inflamed colonic mucosa of wild type *RUNX3* patients showed upregulated RUNX3 mRNA expression compared to non-inflamed colonic mucosa of the same patient or controls. In contrast, inflamed ileal mucosa did not show an upregulated RUNX3 mRNA suggesting region-specific upregulation. This corresponds with the association of *RUNX3* SNP rs2236851 with UC and of microsattelite marker RX-02 with CD with colonic localization.
We could not confirm the association of the *SLC22A4/5* -207 G→C and 1672 G→C variants with IBD, CD or UC.[8] The previously described risk haplotype TC was not associated with an increased risk for UC, CD or different phenotypes of CD. Interestingly, we did find an association for homozygosity for two other SNPs rs272893 and rs273900 with ORs of 2.19 and 2.36 respectively for CD compared to controls. This effect was even more obvious in CD patients with an early age of onset, ileocolonic localization, non-stricturing non-penetrating behavior, extra-intestinal manifestations and, although numbers are small, also with upper GI tract localization. This finding supports the evidence that the *SLC22A4/5* variants are potential candidate genes for IBD susceptibility. However, due to the high degree of LD in the IBD5 locus it remains to be seen whether *SLC22A4/5* are responsible for CD susceptibility or that they are in strong LD with other truly causative genes.
The interaction of a member of the *RUNX* family and *SCL22A4* in patients with RA is intriguing.[17] Therefore, we also analyzed the previously described SNP disrupting the RUNX binding

site in *SLC22A4*. We did not find evidence for an association of this particular SNP. with CD or UC susceptibility. However, carriership of one of the risk associated alleles of *RUNX3* in combination with one of the risk associated alleles of *SLC22A4/5* showed an increased risk for having UC compared to CD (OR 3.83). Contrary to previous observations reporting epistasis for the 1p locus, where *RUNX3* resides, and IBD1, the locus of *CARD15*, no interaction could be found of *RUNX3* with *CARD15*.[15] We did not find the previously described epistasis between *SLC22A4/5* and *CARD15*.[8]

Transport proteins such as multidrug resistance protein 1 (MDR1) and OCTN have been implicated in susceptibility to inflammatory bowel disease. Both *MDR1* and *SLC22A4* genes have RUNX binding sites in their promoter regions. An intronic SNP in a RUNX1 binding site of *SLC22A4* resulted in a stronger suppression of OCTN expression.[17] Furthermore, RUNX3 is capable of suppressing MDR-1 [29] We showed a correlation between upregulated RUNX3 and downregulated OCTN in colonic but not in ileal tissue. In addition to mutations in *SLC22A4* and *MDR1*genes, *RUNX* mutations might further decrease tissue specific expression of these susceptibility genes indicating that a gene (*SLC22A4/5*, *MDR1*) with a susceptibility variant for IBD and a regulatory molecule (*RUNX*) with an allelic different binding for this susceptibility variant might further increase the risk of disease.

In summary, we provide evidence for the genetic association of *RUNX3* with ulcerative colitis. Furthermore genetic association for Crohn's Disease with *SLC22A4/5* was confirmed, although this involved two other SNPs than previously described. An epistatic effect of *RUNX3* and *SLC22A4* was found with an increased risk for ulcerative colitis. Our data, combined with the increasingly recognized role of RUNX3 in autoimmunity, suggest an important role for RUNX3 in UC susceptibility. Functional assays for RUNX3 activity are required to clarify the role of *RUNX3* mutations in de development of UC.

Chapter 5

Genetic susceptibility has a more important role in pediatric-onset than in adult-onset Crohn's disease

L. de Ridder, R.K. Weersma, G. Dijkstra, G. van der Steege, M.A. Benninga, I.M. Nolte, J.A.J.M. Taminiau, D.W. Hommes, P.C.F. Stokkers

Inflammatory Bowel Diseases 2007;13:1083-92

Abstract

Background: Genetic susceptibility may have a more important role in the etiology of early-, than of late-onset IBD and may have a higher frequency of gene mutations. We aimed to determine genotype and phenotype in pediatric-onset IBD, to compare these data with adult-onset IBD and controls and to identify genotype-phenotype associations.

Methods: Polymorphisms R702W, G908R and 3020insC of *CARD15*; Asp299Gly and Thr399Ile of *TLR4*; -207G→C, 1672C→T (L503F), rs3792876, rs274551, rs272893 and rs273900 of *SLC22A4/5*; and 113G→A as well as rs2289311, rs1270912 and rs2165047 of *DLG5* were assessed in 103 pediatric-onset and 696 adult-onset IBD patients. Phenotypic classification was based on disease localization and behaviour.

Results: Homozygosity for the 3020Cins mutation was significantly higher in pediatric-onset than in adult-onset CD (4.2% v 0.6%, 95% confidence interval (CI) 1.2-42.0). Homozygosity for the *SLC22A4/5* rs3792876 SNP was significantly higher in pediatric-onset CD than in adult-onset CD (6.1% v 1.1%, p=0.02). The 3020Cins mutation was associated with ileal involvement (1.9% v 13.3%, CI 1.0-53.8) and a positive family history (6.1% v 20%, CI 1.2-9.0). *DLG5* rs2165047 was significantly associated with perianal disease (50% v 21.2%, CI 1.4-4).

Conclusions: 3020Cins in *CARD15* and *SLC22A4/5* rs3792876 mutations occurred statistically significant more often in pediatric-onset compared to adult-onset CD. 3020Cins (*CARD15*) and *DLG5* rs2165047 mutations in this pediatric-onset CD cohort were associated with specific phenotypes.

Introduction

Inflammatory Bowel Diseases (IBD) is a group of disorders marked by chronic inflammation of the intestinal tract. IBD comprehends two distinct disease entities, ulcerative colitis (UC) and Crohn's disease (CD). These two groups are defined on clinical and histo-pathological features, but overlapping syndromes occur. As a multi-factorial disorder IBD is caused by a complex interaction of genetic, bacterial and immunological factors. The disease presents before the age of 20 years in 20-25% of patients. The incidence of pediatric IBD (under the age of 18 years) in the Netherlands is recently assessed and concerns 5.2/100.000 children.[1] This is comparable to data from other European countries.[2 3]

Genetic risk factors for IBD have been extensively studied in the last decades. Through fine mapping of genome wide scans and linkage studies the *CARD15* gene on chromosome 16 has been identified as a susceptibility gene for IBD.[4-6] The polymorphisms R702W, G908R and 3020Cins in the *CARD15* gene are independently associated with an increased risk of susceptibility for Crohn's disease in Caucasian patients. Mutations of the *CARD15* gene are associated with ileal disease activity, the presence of disease at a younger age, more frequent formation of granulomas and the occurrence of penetrating or stricturing phenotypes.[7-11] These can be demonstrated in 5-15% of the Caucasian patients.[12] A high diversity of these mutations exists in different (healthy) European populations. In a mixed European population the 3 mutations R702W, G908R and 3020Cins are found in 4.0, 1.0 and 1.9%, respectively while in a Dutch population the mutations G908R and 3020Cins are found in 3.0 and 1.0%, respectively.[10,13]

The NOD2 proteins and the Toll-like receptors (TLR's) are pattern recognition receptors that signal presence of bacterial antigens, and play a key role in the innate immunity system. NOD2 is located intracellular, whereas TLR's are extracellular proteins. Both NOD2 and the TLR's are activated by pathogen-associated molecular patterns (PAMP) such as endotoxins. Recent studies indicate that members of the TLR family may also increase IBD susceptibility.[14] The Asp299Gly polymorphism in *TLR4* was associated with CD and UC.[15]Another *TLR4* mutation, Thr399Ile was associated with UC.[16]

Of late, 2 novel IBD susceptibility genes have been identified, being *SLC22A4*/5 and *DLG5*. Within the IBD5 locus, a 250 kb risk haplotype on chromosome 5q31 has been associated with CD in different cohorts.[17-19] However, the causative gene has not been clearly recognized due to a high degree of linkage disequilibrium (LD) in this region. Peltekova *et al*. identified 2 functionally relevant mutations in the carnitine/organic cation transporter (OCTN) genes on the IBD5 locus being responsible for the IBD5 association. These mutations concern variants of *SLC22A4* and *SLC22A5* encoding OCTN1 and OCTN2.[20] The T substitution of C1672 in exon 9 of the *SLC22A4* gene encodes for OCTN1 and a G to C substitution at position -207 in the promoter region of *SLC22A5* gene encodes for OCTN2. Together these polymorphisms form a 2 allele risk haplotype associated with CD susceptibility. Interaction of this haplotype with mutations in the *CARD15* gene is suggested.[21] Other studies could not confirm this association of the 2 allele risk haplotype of *SLC22A4*/5 and CD, though in one study an increased risk in a subgroup with perianal CD has been found.[22 23]

Recently Stoll *et al*. identified an association between genetic variations in *DLG5* (Drosophila Discs Large Homologue 5) on chromosome 10q23 and the risk of developing IBD.(24) *DLG5*

is important in maintaining the epithelial structure and genetic variants could result in impaired intestinal permeability. Two haplotypes were involved in IBD and CD susceptibility. A SNP 113G→A, resulting in an amino acid substitution R30Q, was positively associated with IBD and CD patients. A second haplotype identified by 8 other tagging SNPs was protective for IBD. Gene-gene interaction with *CARD15* in CD was detected by a significant difference in association of the 113A variant in patients carrying the risk alleles for *CARD15* compared to patients not carrying these alleles. However several subsequent studies failed to show any association for *DLG5* with IBD.[25-29]

Genetic susceptibility may have a more important role in the etiology of early-, than of late-onset IBD. After all, early-onset IBD patients were less exposed to environmental factors than late-onset IBD patients. If so, a higher frequency of the gene mutations can be expected in pediatric IBD patients. The purpose of our study was to determine the frequency of *CARD15*, *TLR4*, *SLC22A4*/5 and *DLG5* mutations in pediatric-onset IBD, compare these data with adult-onset IBD and to identify genotype-phenotype associations.

Methods

Study patients

From November 1st 2003 to February 1st 2006, 103 pediatric-onset (under the age of 19) IBD patients agreed to participate in this genetic study performed in the Emma Children's Hospital/Academic Medical Center. From January 1st 1997 to April 1st 2005, in the Academic Medical Center, 696 adult-onset (from the age of 19) IBD patients agreed to participate in this study. IBD was diagnosed based on clinical, endoscopic, radiological and histological criteria of Lennard-Jones.[30] Patients with indeterminate colitis were excluded. Of all patients, written informed consent was obtained, and the study protocol was approved by the institutional review board of the Emma Children's Hospital/Academic Medical Center, Amsterdam. Haplotypes of spouses of IBD patients from the University Medical Center Groningen were taken as controls. Adult controls were considered adequate since genetic profiles do not change once a child has become an adult. Most of them were already used for genotype-phenotype studies on *CARD15* and *TLR4*.[14 31]

Phenotype analysis

Gender and age at diagnosis of all patients were assessed. Phenotypic classification was based on disease localization and behaviour. Localization was determined by endoscopic, histological and/or radiological examination. Categories were made according to the Vienna classification: upper gastrointestinal tract (including jejunum or upper ileum), ileal, ileo-colonic (terminal ileum or terminal ileum and ascending colon), colonic (transverse colon, descending colon, sigmoid or rectum) and perianal disease. Behaviour was defined as uncomplicated (non-stricturing/non-penetrating), penetrating or stricturing. Extra-intestinal manifestations included eye, joint, skin and liver involvement. Surgical intervention was defined as any operative IBD-related procedure, such as gut resection or fistula correction. Family history

was defined as positive if at least one first- or second-degree relative was diagnosed with IBD. Phenotypic assessment was mostly available by the IBD-database and completed by the treating pediatric gastroenterologists, the treating gastroenterologists and by chart review.

Genotype analysis

Venous blood samples (10 ml from each pediatric patient and 20 ml from each adult patient) were collected. DNA was extracted following standard procedures and was stored at -80 oC. Primers to amplify the polymorphic loci were selected using on-line Primer3 software. SNP genotyping was carried out by using TaqMan PCR primer/probe sets, designed through Applied Biosystems' Assay by Design service (http://myscience.appliedbiosystems.com/, Foster City, USA). SNP assay reactions were performed in 5 µl volumes and contained 25 ng DNA, 1x TaqMan Universal PCR Master Mix (Applied Biosystems), 100 nM of each primer and 900 nM of each probe. Cycling conditions on the ABI prism 7900 HT (Applied Biosystems) were 2 minutes 50°C, 10 minutes 95°C followed by 40 cycles of 15 seconds 92°C and 1 minute 60°C. End-point fluorescence was measured immediately after cycling. Alleles were assigned using SDS 2.0 software (Applied Biosystems).

DNA of the patients was screened for the R702W, G908R and 3020Cins of the *CARD15* gene and the Asp299Gly and Thr399Ile polymorphisms of *TLR4*. For *SLC22A4/5* the known polymorphisms -207G→C and 1672C→T (L503F) and 4 additional SNPs (rs3792876, rs274551, rs272893 and rs273900) were analyzed. These SNPs were selected as haplotype-tagging SNPs for the most common haplotypes covering the entire *SLC22A4/5* gene region. For *DLG5*, the previously described 113G→A polymorphism was determined. SNP rs2289311 was used as a haplotype-tagging SNP for the undertransmitted *DLG5*_26 haplotype described by Stoll *et al*.(24) Two added SNPs (rs1270912 and rs2165047) were selected as haplotype-tagging SNPs to cover *DLG5* based on information from the HAPMAP database (http://www.hapmap.org).

Correlations

Genotype correlations between pediatric-onset IBD, adult-onset IBD and healthy control populations were examined. Phenotype correlations in pediatric-onset and adult-onset IBD population were examined. Consequently, genotype-phenotype correlations in pediatric-onset IBD were examined and compared to genotype-phenotype correlations in adult-onset IBD.

Statistics

The differences between the frequencies of the *CARD15*, *TLR4*, *SLC22A4*/5 and *DLG5* polymorphisms in pediatric-onset patients were compared with adult-onset patients and healthy controls using χ^2, when valid, otherwise using Fisher's exact tests. Genotypic association analysis of CD characteristics was also performed in subsets of CD patients stratified according to the Vienna classification and comparing them with controls. Gene-gene interactions (*CARD15* interaction with *TLR4*, *SLC22A4*/5 and *DLG5*) were tested by means of logistic regression in SPSS. A p-value of <0.05 was considered to be significant.

Table 1a. IBD patient characteristics.

	Pediatric-onset IBD (n=103)	**Adult-onset IBD (n=612)**
Gender (male; female), n (%)	50; 53 (49; 51)	263; 349 (43.0; 57.0)
Age at diagnosis, mean (range in years)	12.0 (0.5-18)	35.3 (19-69)
Type of IBD (CD; UC), n (%)	72; 31 (69.9; 30.1)	386; 226 (63.1; 36.9)
Ethnicity (caucasian; negroid; asiatic), n (%)	88; 9; 6 (85.4; 8.7; 5.8)	NA
Family history of IBD, n (%)	18 (17.5)	NA
Localization UC (proctitis, left-sided, pancolitis), n (%)	6; 6; 19 (19.4; 19.4; 61.3) [#]	60; 112; 54 (26.5; 49.6; 23.9)

Table 1b. CD patient characteristics.

	Pediatric-onset CD (n=72)	**Adult-onset CD (n=368)**
Gender (male; female), n (%)	40; 32 (56; 44)	134; 252 (35; 65)
Age at diagnosis, mean (range in years)	14 (6-18)	42 (19-70)
Localization (ileal; colonic; ileo-colonic), n (%)	16; 26; 29 (22.2; 36.1; 40.3)[$]	153; 100; 76; (41.6; 25.9; 19.7)
Localization upper; peri-anal, n (%)	NA; 13 (NA; 18)	176; NA (45.6; NA)
Disease behaviour, n (%) (nonstricturing/ nonpenetrating; structuring; penetrating)	51, 10, 14 (70.8; 13.9; 19.4)	176; 81; 129 (45.6; 21.0; 33.4)
Extraintestinal manifestations, n (%)	14 (19.4)	97 (25.1)
Operated patients, n (%)	17 (23.6)	173 (44.8)

[#] $p<0.0001$
[$] $p=0.0004$
Abbreviations: IBD= Inflammatory Bowel Diseases; CD= Crohn's disease; UC= ulcerative colitis

Results

Phenotypes

A total of 103 pediatric-onset IBD patients and 696 adult-onset IBD patients were included. Patient characteristics are described in Table 1a and 1b.

Isolated ileal disease is more common in adult-onset CD when compared to pediatric-onset CD ($p=0.0004$). Pediatric-onset UC more often is a pancolitis when compared to adult disease ($p<0.0001$). Adult-onset CD patients more often have stricturing and/or penetrating disease behaviour and more frequently undergo surgery in comparison to pediatric-onset CD patients.

Table 2a. Genotype and allele frequencies of *CARD15* mutations in pediatric-onset IBD patients, adult-onset CD patients and controls.

		Pediatric-onset IBD	Pediatric-onset CD	Pediatric-onset UC	Adult-onset CD	Controls
R702W						
Genotype, n (%)						
	Wild-type	81 (79.4%)	58 (80.6%)	23 (76.7%)	289 (82.3%)	254 (93.4%)
	Homozygous	2 (2.0%)	2 (2.8%)	0 (0.0%)	7 (2.1%)	1 (0.4%)
	Heterozygous	19 (18.6%)	12 (16.7%)	7 (23.3%)	51 (15.5%)	17 (6.3%)
p-value (v controls)		0.0003	0.003	0.006		
p-value (v adult-onset)			NS			
p-value (adult-onset v controls)					0.0006	
Mutant allele, n (%)		23 (11.3%)	16 (22.2%)	7 (11.7%)	65 (9.4%)	19 (3.5%)
G908R						
Genotype, n (%)						
	Wild-type	97 (94.2%)	66 (91.7%)	31 (100%)	289 (90.0%)	256 (94.1%)
	Homozygous	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.9%)	2 (0.7%)
	Heterozygous	6 (5.8%)	6 (8.3%)	0 (0.0%)	29 (9.0%)	18 (6.5%)
p-value (v controls)		NS	NS	NS		
p-value (v adult-onset)			NS			
p-value (adult-onset v controls)					NS	
Mutant allele, n (%)		6 (2.9%)	6 (4.2%)	0 (0.0%)	35 (5.5%)	22 (4.0%)
3020insC						
Genotype, n (%)						
	Wild-type	91 (88.3%)	62 (86.1)	29 (93.5%)	302 (88.0%)	252 (92.6%)
	Homozygous	3 (2.9%)	3 (4.2%)	0 (0.0%)	2 (0.6%)	2 (0.7%)
	Heterozygous	9 (8.7%)	7 (9.7%)	2 (6.5%)	39 (11.4%)	18 (6.6%)
p-value (v controls)		NS	NS	NS		
p-value (v adult-onset)			NS			
p-value (adult-onset v controls)					NS	
Mutant allele, n (%)		15 (7.3%)	13 (9.0%)	2 (3.2%)	43 (6.3%)	22 (4.0%)
CARD15	Compound heterozygous and/or homozygous	9 (8.7%)	8 (11.1%)	1 (3.2%)	21 (7.2%)	8 (2.9%)
p-value (vs controls)		0.017	0.005	NS		
p-value (vs adult-onset)			NS			
p-value (adult-onset v controls)					0.006	

Table 2b. Genotype and allele frequencies of *TLR4* mutations in pediatric-onset IBD patients, adult-onset IBD patients and controls.

		Pediatric-onset IBD	Pediatric-onset CD	Pediatric-onset UC	Adult-onset IBD	Adult-onset CD	Adult-onset UC	Controls
Asp299Gly								
Genotype, n (%)								
	Wild-type	89 (86.4%)	62 (86.1%)	27 (87.1%)	515 (85.3%)	320 (84.7%)	195 (86.3%)	224 (91.8%)
	Homozygous	1 (1.0%)	1 (1.4%)	0 (0.0%)	7 (1.2%)	5 (1.3%)	2 (0.9%)	0 (0.0%)
	Heterozygous	13 (12.6%)	9 (12.5%)	4 (12.9%)	82 (13.6%)	53 (14.0%)	29 (12.8%)	20 (8.2%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					0.02	0.02	NS	
Mutant allele, n (%)		15 (7.3%)	11 (7.6%)	4 (6.5%)	96 (7.9%)	63 (8.3%)	33 (7.3%)	20 (4.1%)
Thr399Ile								
Genotype, n (%)								
	Wild-type	90 (87.4%)	63 (87.5%)	27 (87.1%)	486 (84.5%)	298 (83.2%)	188 (86.6%)	224 (91.1%)
	Homozygous	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (0.7%)	3 (0.8%)	1 (0.5%)	0 (0.0%)
	Heterozygous	13 (12.6%)	9 (12.5%)	4 (12.9%)	85 (15.2%)	57 (15.9%)	28 (12.9%)	22 (8.9%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					0.03	0.01	NS	
Mutant allele, n (%)		13 (6.3%)	9 (6.3%)	4 (6.5%)	93 (8.3%)	63 (8.8%)	30 (6.9%)	22 (4.5%)
TLR4	Compound heterozygous and/or homozygous	13 (12.6%)	9 (12.5%)	3 (9.7%)	73 (13.1%)	47 (13.7%)	26 (12.6%)	18 (7.4%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (CD adult-onset v controls)					0.02	0.02	NS	

Association analyses

Genotype and allele frequencies for the *CARD15* polymorphisms R702W, G908R and 3020Cins are detailed in table 2a.

A statistical association for R702W was found with pediatric-onset CD (carriership in patients with CD 14 [19.5%] and in patients with pediatric-onset UC 7 [23.3%] versus controls 18 [6.6%]: CD p=0.0027, UC p=0.0063). For the 3020Cins SNP, a nearly statistical association was found with pediatric-onset CD (carriership of 3020Cins in patients with CD 10 [13.9%] versus controls 20 [7.3%]: p=0.060). Homozygosity for the 3020Cins mutation was noted in 3 patients (4.2%) of the pediatric-onset CD cohort, whereas 2 out of 343 patients (0.7%) of the adult-onset CD patients were homozygous (p=0.04, relative risk [RR] 7.1 95% confidence interval [CI] 1.2-42.0). The prevalence of *CARD15* mutant homozygotes and compound heterozygotes was higher in the pediatric cohort compared to adult-onset CD, but this did not reach statistical significance: 8 out of 72 (11.1%) pediatric-onset CD patients and 21 out of 293 adult-onset CD patients (7.2%) were either compound heterozygous or homozygous for *CARD15* mutations. It was however significantly increased compared to controls (8 [2.9%], p=0.005). The polymorphism R702W was significantly increased in adult-onset CD patients compared to controls (p=0.0006). The prevalence of *CARD15* mutant homozygotes and compound heterozygotes also was increased in adult-onset CD patients versus controls (p=0.006, RR1.6 95% CI 1.2-2.0).

For the Asp299Gly and Thr399Ile polymorphisms of *TLR4* no significant associations were found between pediatric-onset IBD, CD or UC versus controls or versus adult-onset IBD, CD or UC patients (table 2b).

However, both polymorphisms Asp299Gly and Thr399Ile were significantly increased in adult-onset CD patients compared to controls (p=0.02 and p=0.01, respectively). The prevalence of *TLR4* mutant homozygotes and compound heterozygotes also was increased in adult-onset CD patients versus controls (p=0.02; RR 1.28; 95% CI 1.1-1.5). The *CARD15* and TLR Asp299Gly data of the adult-onset IBD patients partially overlap with the cohort analysed and described by Braat *et al*.[32]

For the *SLC22A4/5* rs3792876 SNP a statistical association was found with pediatric-onset CD (p=0.009 versus controls and p=0.03 versus adult-onset CD) (table 2c).

Homozygosity for the *SLC22A4/5* rs3792876 SNP also appeared significantly increased in pediatric-onset CD patients when compared to the adult-onset CD patients (4 [6.1%] versus 4 [1.1%]; p=0.02). No association with *SLC22A4/5* was found for the 2 previously described SNPs L503F and -207G→C and 3 other tagging SNPs with IBD, CD or UC. The *SLC22A4/5* rs272893 and rs273900 SNPs were significantly increased in adult-onset CD patients compared to controls (p=0.04). The *SLC22A4/5* rs274551 SNP was significantly decreased in adult-onset CD patients compared to controls (p=0.002). Ninety-three pediatric-onset IBD patients were genotyped for the *DLG5* polymorphisms (table 2d).

Carriership of mutations was equally divided between men and women. For *DLG5*, rs2165047 tended to be differently distributed among pediatric-onset UC compared to controls, although this was not statistically significant (p=0.088). No associations were found versus the adult-onset IBD population. In adult-onset CD and UC however, the different distribution of *DLG5* rs2165047 did reach statistical significance compared to controls (p=0.04 and p=0.03 respectively).

Table 2c. Genotype and allele frequencies of *SLC22A4/5* mutations in pediatric-onset IBD patients, adult-onset IBD patients and controls

		Pediatric-onset IBD	Pediatric-onset CD	Pediatric-onset UC	Adult-onset IBD	Adult-onset CD	Adult-onset UC	Controls
rs3792876								
Genotype, n (%)								
	Wild-type	79 (85.0%)	54 (81.9%)	24 (96.0%)	570 (83.2%)	315 (84.4%)	255 (83.9%)	248 (86.7%)
	Homozygous	4 (4.3%)	4 (6.1%)	0 (0.0%)	5 (0.7%)	4 (1.1%)	1 (0.3%)	1 (0.4%)
	Heterozygous	10 (10.8%)	8 (12.1%)	1 (4.0%)	110 (16.1%)	62 (14.5%)	48 (15.8%)	37 (12.9%)
p-value (v controls)		0.03	0.009	NS				
p-value (v adult-onset)		0.01	0.03	NS				
p-value (adult-onset v controls)					NS	NS	NS	
Mutant allele, n (%)		18 (9.7%)	16 (12.1%)	1 (2.0%)	112 (8.2%)	62 (8.3%)	50 (8.2%)	39 (6.8%)
p-value (v controls)		NS	0.04	NS				
rs272893								
Genotype, n (%)								
	Wild-type	44 (47.8%)	28 (43.8%)	14 (53.8%)	274 (40.5%)	158 (42.5%)	116 (38.0%)	124 (42.6%)
	Homozygous	10 (10.9%)	8 (12.5%)	2 (7.7%)	92 (13.6%)	55 (14.8%)	37 (12.1%)	25 (8.8%)
	Heterozygous	38 (41.3%)	28 (43.8%)	10 (38.5%)	311 (45.9%)	159 (42.7%)	152 (49.8%)	141 (48.6%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					NS	0.04	NS	
Mutant allele, n (%)		58 (31.5%)	44 (34.3%)	14 (26.9%)	495 (36.6%)	269 (36.2%)	226 (37.0%)	191 (32.9%)
rs1050152								
Genotype, n (%)								
	Wild-type	31 (33.7%)	24 (36.9%)	7 (28.0%)	225 (33.1%)	117 (31.1%)	108 (35.5%)	88 (30.0 %)
	Homozygous	19 (20.7%)	14 (21.5%)	4 (16.0%)	135 (19.9%)	76 (20.2%)	59 (19.4%)	55 (18.8%)
	Heterozygous	42 (45.7%)	38 (41.5%)	14 (56.0%)	320 (47.1%)	183 (48.7%)	137 (45.1%)	150 (51.2%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					NS	NS	NS	
Mutant allele, n (%)		80 (43.5%)	66 (50.8%)	22 (44%)	590 (43.4%)	335 (44.5%)	255 (41.9%)	260 (44.5%)

Table 2c. Continued

rs273900								
Genotype, n (%)								
	Wild-type	36 (40.4%)	23 (37.1%)	12 (46.2%)	269 (38.6%)	154 (39.4%)	115 (37.7%)	125 (43.1%)
	Homozygous	10 (11.2%)	8 (12.9%)	2 (7.7%)	100 (14.4%)	63 (16.1%)	37 (12.1%)	24 (8.3%)
	Heterozygous	43 (48.3%)	31 (50.0%)	12 (46.2)%	327 (47.0%)	174 (44.5%)	153 (50.2%)	141 (48.6%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					NS	0.01	NS	
Mutant allele, n (%)		63 (35.4%)	47 (37.9%)	16 (30.8%)	527 (75.8%)	300 (38.4%)	227 (37.2%)	189 (32.6%)
rs274551								
Genotype, n (%)								
	Wild-type	58 (63.0%)	43 (65.2%)	14 (56.0%)	481 (70.8%)	276 (74.0%)	205 (67.0%)	187 (64.8%)
	Homozygous	5 (5.4%)	4 (6.1%)	1 (4.0%)	17 (2.5%)	6 (1.6%)	11 (3.6%)	18 (6.3%)
	Heterozygous	29 (31.5%)	19 (28.8%)	10 (40.0%)	181 (26.7%)	91 (24.4%)	90 (29.5%)	83 (29.0%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					0.01	0.002	NS	
Mutant allele, n (%)		39 (21.2%)	27 (20.5%)	12 (24.0%)	215 (15.8%)	103 (13.8%)	112 (18.3%)	119 (20.7%)
rs2631367								
Genotype, n (%)								
	Wild-type	18 (19.8%)	14 (21.9%)	4 (16.0%)	173 (26.9%)	94 (26.8%)	79 (27.0%)	70 (25.0%)
	Homozygous	22 (24.2%)	18 (28.1%)	3 (12.0%)	147 (22.8%)	79 (22.5%)	68 (23.2%)	66 (23.4%)
	Heterozygous	51 (56.0%)	32 (50.0%)	18 (72.0%)	324 (50.3%)	178 (50.7%)	146 (49.8%)	145 (51,6%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					NS	NS	NS	
Mutant allele, n (%)		95 (52.2%)	68 (53.1%)	24 (48.0%)	618 (48.0%)	336 (47.9%)	282 (48.1%)	277 (49.3%)

Table 2d. Genotype and allele frequencies of *DLG5* mutations in pediatric-onset IBD patients, adult-onset IBD patients and controls.

		Pediatric-onset IBD	Pediatric-onset CD	Pediatric-onset UC	Adult-onset IBD	Adult-onset CD	Adult-onset UC	Controls
rs2165047								
Genotype, n (%)								
	Wild-type	45 (48.4%)	35 (53.9%)	9 (34.6%)	331 (49.0%)	181 (48.8%)	150 (49.2%)	161 (55.7%)
	Homozygous	8 (8.6%)	6 (9.2%)	2 (7.7%)	67 (9.9%)	35 (9.4%)	32 (10.5%)	14 (4.9%)
	Heterozygous	40 (43.1%)	24 (36.9%)	15 (57.7%)	278 (41.1%)	155 (41.8%)	123 (40.3%)	114 (39.4%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					0.02	0.04	0.03	
Mutant allele, n (%)		56 (30.1%)	36 (27.7%)	19 (21.2%)	412 (30.5%)	225 (30.3%)	187 (30.7%)	142 (24.6%)
rs2289311								
Genotype, n (%)								
	Wild-type	42 (45.7%)	29 (44.6%)	13 (52.0%)	333 (49.5%)	175 (47.4%)	158 (52.0%)	120 (41.3%)
	Homozygous	8 (8.7%)	5 (7.7%)	3 (12.0%)	75 (11.1%)	39 (10.6%)	36 (11.8%)	43 (14.9%)
	Heterozygous	42 (45.7%)	31 (47.7%)	9 (36%)	265 (39.4%)	155 (42.0%)	110 (36.2%)	126 (43.8%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					NS	NS	NS	
Mutant allele, n (%)		58 (31.5%)	41 (31.5%)	15 (30%)	415 (30.8%)	233 (31.6%)	182 (29.9%)	212 (36.7%)
rs1270912								
Genotype, n (%)								
	Wild-type	43 (46.7%)	31 (47.0%)	11 (44.0%)	316 (46.7%)	174 (46.9%)	142 (46.6%)	133 (45.2%)
	Homozygous	12 (13.0%)	8 (12.1%)	4 (16.0%)	84 (12.4%)	44 (11.9%)	40 (13.1%)	31 (10.8%)
	Heterozygous	37 (40.2%)	27 (40.9%)	10 (40.0%)	276 (40.8%)	153 (41.2%)	123 (40.3%)	129 (44.0%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					NS	NS	NS	
Mutant allele, n (%)		61 (33.2%)	43 (32.6%)	18 (36%)	444 (32.8%)	241 (32.5%)	203 (33.3%)	191 (32.6)
rs1248696								
Genotype, n (%)								
	Wild-type	68 (80.4%)	53 (80.3%)	20 (80.0%)	541 (80.1%)	296 (79.1%)	245 (81.4%)	227 (77.6%)
	Homozygous	1 (1.1%)	0 (0.0%)	1 (4.0%)	9 (1.3%)	5 (1.3%)	4 (1.3%)	5 (1.7%)
	Heterozygous	17 (18.5%)	13 (19.7%)	4 (16.0%)	125 (18.5%)	73 (19.5%)	52 (17.3%)	61 (20.6%)
p-value (v controls)		NS	NS	NS				
p-value (v adult-onset)		NS	NS	NS				
p-value (adult-onset v controls)					NS	NS	NS	
Mutant allele, n (%)		19 (10.3%)	13 (9.8%)	6 (12%)	143 (10.6%)	83 (11.1%)	60 (10.0%)	71 (12.1%)

Table 3. Mutant allele frequencies in subgroups of pediatric-onset Crohn's Disease.

Number (%), p values included in case of significance							
	R702W	G908R	3020Cins	Asp299Gly	Thr399Ile	OCTN rs3792876	DLG5 rs2165047
Localisation							
Ileum (n=45)	12 (13.3)	4 (4.4)	12 (13.3)	5 (5,6)	5 (5,6)	9 (10.0)	23 (52.3)
Non ileal (n=27)	4 (7.4)	2 (3.7)	1 (1.9)[$]	4 (7.4)	4 (7.4)	10 (18.5)	19 (35.2)
Behaviour							
Nonstricturing, nonpenetrating (n=51)	15 (14.7)	4 (3.9)	6 (5.9)	6 (5.9)	6 (5.9)	12 (11.8)	24 (23.5)
Stricturing (n=10)	1 (5.0)	1 (5.0)	5 (25.0)[#]	2 (10.0)	2 (10.0)	4 (20)	4 (20.0)
Penetrating (n=14)	2 (7.1)	1 (3.6)	2 (7.1)	3 (10.7)	3 (10.7)	3 (10.7)	11 (39.3)
Operated (n=17)	2 (5.9)	1 (2.9)	6 (17.6)	6 (17.6)	6 (17.6)	6 (17.6)	11 (32.4)
Not operated (n=55)	14 (12,7)	5 (4.5)	7 (6.4)	12 (10.9)	12 (10.9)	11 (10.0)	30 (27.3)
Perianal disease (n=13)	2 (7.7)	1 (3.8)	2 (7.7)	2 (7.7)	2 (7.7)	5 (19.2)	13 (50.0)[*]
No perianal disease (n=59)	14 (11.9)	5 (4.2)	11 (9.3)	7 (5.9)	7 (5.9)	12 (10.2)	25 (21.2)
Extra-intestinal disease (n=14)	7 (25)	1 (3.6)	1 (3.6)	3 (10.7)	3 (10.7)	3 (10.7)	6 (21.4)
No extra-intestinal disease (n=58)	11 (9.5)	5 (4.3)	5 (4.3)	6 (5.2)	6 (5.2)	10 (8.6)	29 (25.0)
Family history of IBD (n=15)	2 (6.7	1 (3.3)	6 (20)[§]	2 (6.7)	2 (6.7)	5 (16.7)	8 (26.7)
No family history of IBD (n=57)	14 (12.3)	5 (4.4)	7 (6.1)	7 (6.1)	7 (6.1)	10 (8.8)	30 (27.3)

[$] p= 0.03 RR 7.2 95.0% CI: 1.0-53.8
[#] p= 0.02 RR 4.1, 95.0% CI: 1.5-11.2 (stricturing versus other behaviour)
[*] p=0.003 RR 2.4 95.0% CI: 1.4-4.0
[§] p=0.07 RR 3.3 95% CI 1.2-9.0

Gene-gene interactions

No interactions of *TLR4*, *SLC22A4*/5 or *DLG5* with *CARD15* were detected.

Genotype-phenotype correlations pediatric-onset CD patients

Genotype-phenotype correlations of the pediatric-onset CD patients are shown in table 3.
The 3020Cins mutation was associated with ileal involvement (including ileocolonic localization) (p=0.03; RR 7.2; 95% CI 1.0-53.8) and this was even more strongly associated with purely ileal disease (p<0.0001; RR 12.2; 95% CI 4.1-36.4). The 3020Cins mutation was suggestively associated with a positive family history (p=0.07).
DLG5 rs2165047 was significantly associated with perianal disease (p=0.003; RR 2.4; 95% CI 1.4-4.0). Other genotype/phenotype correlations were not found in the pediatric-onset CD cohort.

Discussion

This is the first study to report the contribution of the *CARD15, TLR4*, *SLC22A4/5* and *DLG5* genes in a pediatric-onset IBD population and compare these data to an adult-onset IBD population.

CARD15

Homozygosity for the 3020Cins NOD2 mutation was significantly more frequent in pediatric-onset CD than in adult-onset CD. This finding confirms our hypothesis that this *CARD15* mutation predisposes for pediatric-onset CD. So far, reports in literature are conflicting. Ferraris *et al*. found a higher incidence of the 3 major *CARD15* mutations in their Italian pediatric vs. adult CD cohort (p=0.056).[33] Weiss *et al*. demonstrated a higher prevalence of G908R mutation in 67 Jewish pediatric vs. adult CD patients, though this did not reach statistical significance (p=0.063).[34] Others did not find any differences in the prevalence of the 3 major *CARD15* mutations between a pediatric-onset and an adult-onset CD cohort.[35 36]
These conflicting results can be explained by large regional and ethnical differences in genotypes, the broad spectrum of phenotypes within IBD and the relatively low numbers of patients included in these studies. A statistical association between R702W and a nearly statistical association between 3020Cins and CD was found. The genotype-phenotype analysis revealed a strong association between the 3020Cins mutation and CD localization in the terminal ileum. The association between *CARD15* mutations and involvement of the ileum is well described, both in adult and pediatric populations.[8 37-39] Furthermore, we found that the 3020Cins mutation is suggestively associated with familial disease. This finding is confirmed by a meta-analysis of 42 studies (mainly adult patients), in which more familial disease in *CARD15* mutation carriers was demonstrated.[39]

TLR4

No significant associations were found for the *TLR4* polymorphisms in this pediatric-onset IBD cohort, comparable with the findings of Leshinsky-Silver *et al*.[36] However, both Asp299Gly

and Thr399Ile polymorphisms were associated with CD in our adult-onset cohort. Moreover, these polymorphisms have been associated with CD and UC.[15 40-42] In other studies, these associations were not confirmed.[43 44]

SLC22A4/5

In our pediatric-onset CD cohort we cannot confirm the associations of the 2 variants in *SLC22A4* and *SLC22A5* that were reported by Peltekova *et al*. They found that the association with the IBD5 locus is strongest for CD patients diagnosed under 16 years of age and this finding was confirmed in another study.[17 20] Nevertheless, Russell *et al*. reported that the *SLC22A4/5* variants do not act independently of variants in the IBD5 locus.[45] In our cohort, although numbers are small, SNP rs3792876 was statistically associated with pediatric-onset CD compared with healthy controls. This mutation was also more frequent in the pediatric-onset CD patients when compared to the adult-onset CD patients. As far as we know, this finding has not been reported before. Taken together, these data support the evidence that the *SLC22A4/5* variants are genetic risk factors for CD susceptibility. However, due to the high degree of LD in the IBD5 locus it remains to be seen whether *SLC22A4/5* variants are responsible for CD susceptibility or that they are in strong LD with other true causative genetic variants.
Moreover, we observed a statistically significant higher frequency of the *SLC22A4*/5 rs3792876 mutation in the pediatric-onset CD patients compared to the adult-onset CD patients. As far as we know, this comparison is not reported before.

DLG5

We could not confirm the previously described association of *DLG5* and IBD in our pediatric cohort. Interestingly, we did find that rs2165047 was significantly increased in both our adult-onset CD and UC population. Several other studies failed to validate the significance of *DLG5* variants as important determinants in IBD susceptibility in adult-onset IBD.[25 27-29] Friedrichs *et al*. recently demonstrated that carriership of *DLG5* mutations was a susceptibility factor for CD in men but not in women as did Biank *et al*. in pediatric-onset CD cohort.[46 47] However, *DLG5* mutation carriership was equally divided among both sexes in both our pediatric-onset CD patients and the healthy control group. Associations with *DLG5* mutation carriership were not found.
DLG5 rs2165047 mutation carriership in our pediatric CD cohort was significantly associated with perianal disease. Larger studies are needed to confirm these data since the number of patients of this cohort is too small to infer any conclusions as yet. We don't have data on ethnicity of the adult-onset population. However, it is expected to be the same in the pediatric-onset as in the adult-onset population, as it concerns the same geographic area where the patients come from and referral is based on the same conditions. Calculations were also performed in an exclusively Caucasian pediatric-onset IBD cohort and no other statistical significances were revealed compared with the entire cohort, although some became stronger (data not shown).

In conclusion, this study demonstrated that the R702W mutation in *CARD15* is associated with CD in a Dutch pediatric-onset IBD cohort. Secondly, for *SLC22A4/5*, SNP rs3792876 is

associated with pediatric-onset CD. Moreover, both 3020Cins and *SLC22A4/5* rs3792876 mutations occurred statistically significant more often in pediatric-onset compared to adult-onset CD. Since *TLR4*, *SLC22A4/5* and *DLG5* seem moderate risk alleles, more associations with pediatric-onset IBD cannot be excluded yet and larger pediatric-onset CD cohorts need to be examined. Genetic susceptibility has a more important role in the etiology of early- than of late-onset CD. Within pediatric-onset CD specific genotype-phenotype associations can be found. These data stress the importance of genetic susceptibility research in large pediatric-onset IBD cohorts in order to find new genes and to establish the influence of these mutations on disease behaviour.

References

1. Zaag-Loonen van der HJ, Casparie M, Taminiau JAJM, Escher JC, Pereira RR, Derkx HH. The incidence of pediatric inflammatory bowel disease in the Netherlands: 1999-2001. J Pediatr Gastroenterol Nutr 2004;38:302-7.

2. Ahmed M, Davies IH, Hood K, Jenkins HR. Incidence of pediatric inflammatory bowel disease in South Wales. Arch Dis Child 2006;91:344-5.

3. Sawczenko A, Sandhu BK, Logan RFA *et al*. Prospective survey of childhood inflammatory bowel disease in the British isles. Lancet 2001;357:1093-4.

4. Hampe J, Cuthbert A, Croucher PJP *et al*. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. Lancet 2003;362:1925-8.

5. Hugot JP, Chamaillard M, Zouali H *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 2001;411:599-603.

6. Ogura Y, Bonen DK, Inohora N *et al*.. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 2001;411:603-6.

7. Ahmad T, Armuzzi A, Bunce M *et al*. The molecular classification of the clinical manifestations of Crohn's disease. Gastroenterology 2002;122:854-66.

8. Cuthbert AP, Fisher SA, Mirza MM *et al*. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease.Gastroenterology 2002;122:867-74.

9. Hampe J, Grebe J, Nikolaus S, Solberg C *et al*. Association of NOD2 (CARD 15) genotype with clinical course of Crohn's disease: a cohort study. Lancet 2002; 359:1661-5.

10. Lesage S, Zouali H, Cezard JP *et al*. *CARD15*/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease. Am J Hum Genet 2002;70:845-57.

11. Radlmayr M, Torok HP, Martin K, Folwaczny C. The c-insertion mutation of the NOD2 gene is associated with fistulizing and fibrostenotic phenotypes in Crohn's disease. Gastroenterology 2002;122:2091-2.

12. Gaya DR, Russell RK, Nimmo ER, Satsangi J. New genes in inflammatory bowel disease: lessons for complex diseases? Lancet 2006;367:1271-84.

13. Murillo L, Crusius JB, van Bodegraven AA, Alizadeh BZ, Pena AS. *CARD15* gene and the classification of Crohn's disease. Immunogenetics 2002;54:59-61.

14. Oostenbrug LE, Drenth JPH, de Jong DJ *et al*. Association between Toll-like receptor 4 and Inflammatory Bowel Disease. Inflamm Bowel Dis 2005;11:567-75.

15. Franchimont D, Vermeire S, El Housni H *et al*. Deficient host-bacteria interactions in inflammatory

bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis.Gut 2004;53:987-92.

16. Török H-P, Glas J, Tonenchi L, Mussack T, Folwaczny C. Polymorphisms of the lipopolysaccharide-signaling complex in inflammatory bowel disease: association of a mutation in the Toll-like receptor 4 gene with ulcerative colitis. Clin Immunol 2004;112:85-91.

17. Rioux JD, Silverberg MS, Daly MJ *et al*. Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. Am J Hum Genet 2000;66:1863-70.

18. Rioux JD, Daly MJ, Silverberg MS *et al*. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genetic 2001;29:223-8.

19. Armuzzi A, Ahmad T, Ling KL *et al*. Genotype-phenotype analysis of the Crohn's disease susceptibility haplotype on chromosome 5q31. Gut 2003;52:1133-9.

20. Peltekova VD, Wintle RF, Rubin LA *et al*. Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet 2004;36:471-5.

21. Newman B, Gu X, Wintle R *et al*. A risk haplotype in the Solute Carrier Family 22A4/22A5 gene cluster influences phenotypic expression of Crohn's disease. Gastroenterology 2005;128:260-9.

22. Vermeire S, Pierik M, Hlavaty T *et al*. Association of organic cation transporter risk haplotype with perianal penetrating Crohn's disease but not with susceptibility to IBD. Gastroenterology 2005;129:1845-53.

23. Noble CL, Nimmo ER, Drummond H *et al*. The contribution of OCTN1/2 variants within the IBD5 locus to disease susceptibility and severity in Crohn's disease. Gastroenterology 2005;129:1854-64.

24. Stoll M, Corneliussen B, Costello CM *et al*. Genetic variation in *DLG5* is associated with inflammatory bowel disease. Nat Genet 2004;36:476-80.

25. Noble CL, Nimmo ER, Drummond H *et al*. *DLG5* variants do not influence susceptibility to inflammatory bowel disease in the Scottish population.Gut 2005;54:1416-20.

26. Török H-P, Glas J, Tonenchi L *et al*. Polymorphisms in the *DLG5* and OCTN cation transporter genes in Crohn's disease. Inflamm Bowel Dis 2005;54:1421-7.

27. Buning C, Geerdts L, Fiedler T *et al*. *DLG5* variants in inflammatory bowel disease. Am J Gastroenterol 2006;101:786-92.

28. Tremelling M, Waller S, Bredin F, Greenfield S, Parkers M. Genetic variants in TNF-alpha but not *DLG5* are associated with inflammatory bowel disease in a large United Kingdom cohort. Inflamm Bowel Dis 2006;12:178-84.

29. Yamazaki K, Takazoe M, Tanaka T *et al*. Association analysis of *SLC22A4*, *SLC22A5* and *DLG5* in Japanese patients with Crohn disease. J Hum Genet 2004;49:664-8.

30. Lennard-Jones JE. Classification of inflammatory bowel disease. Scand J Gastroenterol Suppl 1989;170:2-6.

31. Oostenbrug LE, Nolte IM, Oosterom E *et al*. *CARD15* in inflammatory bowel disease and Crohn's disease phenotypes: An association study and pooled analysis. Dig Liver Dis 2006;38:834-45

32. Braat H, Stokkers P, Hommes DW *et al*. Consequence of functional Nod2 and *TLR4* mutations on gene transcription in Crohn's disease patients. J Mol Med 2005;83:601-9.

33. Ferraris A, Knafelz D, Torres B *et al*. Analysis of *CARD15* gene variants in Italian pediatric patients with inflammatory bowel diseases. J Pediatr 2005;147:272-3.

34. Weiss B, Shamir R, Bujanover Y *et al*. NOD2/*CARD15* mutation analysis and genotype-phenotype correlation in Jewish pediatric patients compared with adults with Crohn's disease. J Pediatr 2004;145:208-12.

35. Tomer G, Ceballos C, Concepcion E, Benkov KJ. NOD2/*CARD15* variants are associated with lower weight at diagnosis in children with Crohn's disease. Am J Gastroenterol 2003;98:2479-84.

36. Leshinsky-Silver E, Karban A, Buzhakor E *et al*. Is age of onset of Crohn's disease governed by mutations in NOD2/Caspase recruitment domains 15 and Toll-like receptor 4? Evaluation of a pediatric cohort. Pediatr Res 2005;58:499-504.

37. Wine E, Reif SS, Leshinsky-Silver E *et al*. Pediatric Crohn's disease and growth retardation: the role of genotype, phenotype, and disease severity. Pediatrics 2004;114:1281-6.

38. Russell RK, Drummond HE, Nimmo EE *et al*. Genotype-phenotype analysis in childhood-onset Crohn's disease: NOD2/*CARD15* variants consistently predict phenotypic characteristics of severe disease. Inflamm Bowel Dis 2005;11:955-64.

39. Economou M, Trikalinos TA, Loizou KT, Tsianos EV, Ioannidis JP. Differential effects of NOD2 variants on Crohn's disease risk and phenotype in diverse populations: a metaanalysis. Am J Gastroenterol 2004;99:2393-2404.

40. Ouburg S, Mallant-Hent R, Crusius JBA *et al*. The toll-like receptor 4 (*TLR4*) Asp299Gly polymorphism is associated with colonic localisation of Crohns's disease without a major role for the Saccharomyces cerevisiae mannan-LBP-CD14-*TLR4* pathway. Gut 2005;54:439-40.

41. Gazouli M, Mantzaris G, Kotsinas K *et al*. Association between polymorphisms in the Toll-like receptor 4, CD14, and *CARD15*/NOD2 and inflammatory bowel disease in the Greek population. World J Gastroenterol 2005;11:681-5.

42. Török HP, Glas J, Tonenchi L *et al*. Polymorphisms of the lipopolysaccharide-signaling complex in inflammatory bowel disease: association of a mutation in the Toll-like receptor 4 gene with ulcerative colitis. Clin Immunol 2004;112:85-91.

43. Arnott ID, Nimmo ER, Drummond HE *et al*. NOD2/*CARD15*, *TLR4* and CD14 mutations in Scottish and Irish Crohn's disease patients: evidence for genetic heterogeneity within Europe? Genes Immun 2004;5:417-25.

44. Lakatos PL, Lakatos L, Szalay F *et al*. Toll-like receptor 4 and NOD2/*CARD15* mutations in Hungarian patients with Crohn's disease: phenotype-genotype correlations. World J Gastroenterol 2005;11:1489-95.

Chapter 6

ATG16L1 and *IL23R* are associated with Inflammatory Bowel Diseases but not with Celiac Disease in the Netherlands

R.K. Weersma, A. Zhernakova, I.M. Nolte, C. Lefebvre, J.D. Rioux, F. Mulder, H.M. van Dullemen, J.H. Kleibeuker, C. Wijmenga, G. Dijkstra

American Journal of Gastroenterology, accepted for publication

Abstract

Background: Inflammatory bowel diseases (IBD) –Crohn's disease (CD) and ulcerative colitis (UC) – and celiac disease are intestinal inflammatory disorders with a complex genetic background. Recently, two novel genes were found to be associated with IBD susceptibility. One, an uncommon coding variant (rs11209026) in the gene encoding for the interleukin-23 receptor(*IL23R*) conferred strong protection against CD. The other, rs2241880 in the autophagy-related 16-like 1 gene(*ATG16L1*) was associated with CD. We performed a case-control study for association of IBD with *IL23R* and *ATG16L1* in a Dutch cohort. We also looked at the association of *IL23R* and *ATG16L1* with celiac disease.

Methods: 518 Dutch Caucasian IBD patients (311 CD and 207 UC, including 176 trios of patients with both parents), 508 celiac disease patients, and 893 healthy controls were studied for association to the rs11209026 (*IL23R*) and rs2241880 (*ATG16L1*) SNPs.

Results: The rs11209026 SNP in *IL23R* had a protective effect for IBD in case-control analysis (OR=0.19; 95% CI: 0.10–0.37; p=6.6E-09). Both CD (OR=0.14; CI: 0.06-0.37; p=3.9E-07) and UC (OR=0.33; CI: 0.15-0.73; p=1.4E-03) were associated with *IL23R*. For *ATG16L1* the rs2241880 SNP was associated with CD susceptibility (OR=1.36; CI: 1.12-1.66; p=0.0017). The population attributable risk for carrying allele G is 0.24 and 0.19 for homozygosity for allele G in CD. No association was found between *IL23R* or *ATG16L1* and celiac disease.

Conclusions: We confirmed the association of *IL23R* and *ATG16L1* with CD susceptibility and also the association of *IL23R* with UC. We found *IL23R* and *ATG16L1* were not associated with celiac disease susceptibility.

Introduction

Inflammatory bowel diseases (IBD) are a group of chronic inflammatory disorders of the gastrointestinal tract of unknown origin; there are two main types: Crohn's disease (CD) and ulcerative colitis (UC).[1] Celiac disease is another chronic inflammatory disorder of the gastrointestinal tract, although it is not usually included under IBD. The combined prevalence of CD and UC is estimated at 100–200/100,000 in developed countries.[2 3] Concordance rates in twins and siblings suggest that a genetic predisposition contributes to the pathogenesis of IBD. [4] In addition to genetic factors, environmental factors and the enteric microbial flora are involved in the development of IBD. Nevertheless the aetiology remains largely unknown.

In 2001 the *CARD15* gene on chromosome 16 was identified as the first susceptibility gene for CD.[5 6] IBD shows a variation in the individual clinical presentation and outcome that is most likely due to differences in the genetic susceptibility, exposure to environmental factors, the commensal bacteria in the intestine, and the intestinal immune system. Apart from the *CARD15* association, the organic cation transporters, encoded by the genes *SLC22A4* and *SLC22A5,* have been reported to be associated with CD susceptibility.[7] However, this association is not independent from the IBD5 haplotype.[8] Further associations have been found with polymorphisms in the genes encoding for Drosophila discs large homologue 5 *(DLG5),* multi-drug resistance gene (*MDR1),* Toll-like receptor 4 (*TLR4*), Caspase-activating recruitment domain 4 *(CARD4)* and Tumour Necrosis Factor superfamily member 15 (*TNFSF15*). However, none of these associations has been as consistently replicated as the CARD15 association. [9-13]

Recently, two genome-wide association studies were performed in IBD and uncovered two novel genes involved in IBD susceptibility: the first study, by Duerr *et al*., identified the gene encoding a subunit for the interleukin 23 receptor (*IL23R*) on chromosome 1p31 as a susceptibility gene for CD in a population of patients with ileal CD of non-Jewish, European ancestry.[14] An uncommon variant Arg381Gln was shown to have a strong protective effect for CD and other non-coding IL23R variants were independently associated. This association was confirmed in independent cohorts of non-Jewish CD and UC patients, not only for CD but also for UC. Since the proinflammatory cytokine IL23 is increasingly being recognized as involved in gut inflammation, this finding is of great importance.[15] The second study, by Hampe *et al*., found a strong association between CD and the autophagy-related 16-like 1 gene (*ATG16L1*) encoding a protein in the autophagosome pathway, which processes intracellular bacteria.[16] They performed a genome-wide association scan in 735 German patients with CD and 368 controls. Their analysis of the most informative SNPs in an independent cohort of CD patients and controls identified *ATG16L1* as a CD susceptibility gene. Marker rs2241880, a coding SNP (Thr300Ala), was strongly associated with susceptibility for CD. The risk for CD was confined to individuals carrying susceptibility allele G at rs2241880. It was suggested that there might be an epistatic effect with one of the three risk alleles in *CARD15* associated with CD.

Since it is of pivotal importance that genetic associations are confirmed in independent cohorts from different countries, we performed a replication study for the two most strongly associated SNPs in *IL23R* and *ATG16L1* in a cohort of IBD patients from the northern part of

the Netherlands. We were also interested in discovering whether these two genes are more generally involved in other common chronic disorders of the gastrointestinal tract and we therefore included a cohort of celiac disease patients from the Netherlands.

Methods

Subjects

IBD: A cohort of 518 IBD patients (311 CD and 207 UC) from the University Medical Center Groningen was used.[17] The diagnosis of CD or UC was based on accepted clinical, radiologic, endoscopic and histopathologic criteria.[1] For 176 (103 CD and 73 UC) patients genotyping of both parents was also available. The clinical characteristics of patients with CD are given in table 1 and with UC in table 2.

Celiac disease: A cohort of 508 independent celiac patients of Dutch origin was included in the study. The diagnosis of CD was made according to the revised ESPGHAN criteria.[18] In addition, the intestinal biopsies on which the initial diagnoses were based were re-evaluated for all these patients by one experienced pathologist, and only patients with a demonstrable villous atrophy and Marsh III lesions were included in this study.

The controls consisted of 893 healthy volunteers recruited from the University Medical Center Utrecht. Celiac cases and controls have been previously described.[19] All the participants were of European Caucasian descent.

All patients gave informed consent and the study was approved by the ethics review committees of the participating hospitals. All DNA samples and data in this study were handled anonymously.

Genotyping and SNP selection

Genotyping assays for the rs11209026 and rs2241880 SNPs were designed for the Sequenom MassArray iPLEX platform using the Sequenom Assay Design software, version 3.0. The case and control samples, as well as 90 Centre d'Etude du Polymorphisme Humain samples included in the International HapMap project were genotyped by primer extension of multiplex PCR products, followed by a chip-based matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF).[20,21] All genotyping was performed according to the manufacturer's protocol at the Laboratory of Genetics and Genomic Medicine (www.inflammgen.org). The DNA samples were processed in 384 well-plates and each 384-plate with patient and control DNA contained 16 genotyping controls (4 duplicates of 4 CEU DNA). Both SNPs were validated and we obtained >99.9% concordance between our genotype data and the CEU data available from HapMap. Laboratory staff was blinded for disease status of each sample.

Statistical analysis

After genotyping the markers, we tested for Hardy-Weinberg equilibrium by comparing the expected and observed genotypes in 2 x 3 χ^2 tables. Both markers showed no deviation from Hardy-Weinberg equilibrium in controls (p-value >0.05). Differences in allele and genotype

Table 1. Clinical characteristics of patients with Crohn's disease.

Total number	311	
Sex (m/f)	106 / 205	
Age at diagnosis (years)		
Mean (range)	31.3	(6.7 – 73.9)
Median 27.7		
Under 40 yrs	227	(73.0%)
Follow up (years)		
Mean (range)	11.8	(0.4 – 49.4)
Median	8.8	
Disease localization		
Ileal	78	(25.1%)
Colon	67	(21.5%)
Ileocolon	166	(53.4%)
Disease behaviour		
Non-stricturing, non penetrating 101	(32.8%)	
Stricturing	76	(24.4%)
Penetrating	133	(42.8%)
Upper GI tract	24	(7.7%)
Peri-anal	88	(28.3%)
Extra-intestinal manifestations	42	(13.5%)
Family history of IBD	62	(19.9%)
History of surgical intervention	187	(60.1%)

Table 2. Clinical characteristics of patients with ulcerative colitis.

Total number	207	
Sex (m/f)	109 / 98	
Age at diagnosis (years)		
Mean (range)	32.9	(5.6 – 46.0)
Median	31.4	
Under 40 yrs	146	(70.5%)
Follow up (years)		
Mean (range)	14.2	(2.5 – 48.9)
Median	12.5	
Disease localization ¶*		
Proctitis	14	(6.8%)
Left-sided	65	(31.4%)
Pancolitis	99	(47.8%)
Extra-intestinal manifestations	34	(16.4%)
Family history of IBD	31	(15.0%)
History of surgical intervention	46	(22.2%)

¶ Defined as maximum disease extend found at endoscopy during follow up.
* Data was unavailable for 29 (14.3%) patients.

distribution in the cases and controls were tested for significance by the chi-squared (χ^2) test. In the trios we looked for association with the transmission disequilibrium test , using the transmission disequilibrium test phase program of the UNPHASED package.[22] Odds ratios were calculated and confidence intervals were approximated using Woolf's method with Haldane's correction. [23] Binary logistic regression was also used to determine gene-gene interaction effects. A significant threshold for p-values was determined at $p<0.05$.

Results

The *IL23R* rs11209026 SNP and the *ATG16L1* rs2241880 SNP were genotyped in a group of 518 IBD patients (311 CD and 207 UC), 508 celiac disease patients and 893 controls, all of European Caucasian descent.

For the rs11209026 SNP in *IL23R* we observed a significant decrease of the rs11209026*A allele in the IBD group compared to controls (OR=0.19; 95% CI: 0.10–0.37; p= 6.6E-09) (Table 3). The overall genotype distribution of the rs11209026 SNP was also significantly different from the control cohort (p=5.3E-08) (Table 3). We investigated whether the protective effect of the rs11209026*A allele was different in the CD and UC patient groups. Both the CD and UC patients showed a much lower frequency of rs11209026*A allele compared to controls (OR=0.14; CI: 0.06-0.37; p=3.9E-07 for CD and OR=0.33; CI: 0.15-0.73; p=1.4E-03 for UC) and in both groups the genotype distribution differed significantly from the controls (Table 3). Although the association of the rs11209026*A allele with CD was stronger than with UC (OR(CD)=0.14 vs. OR(UC)=0.33), this difference was not significant (data not shown).

For 176 IBD patients we also had the genotypes of both parents, allowing us to perform a transmission disequilibrium test analysis. Due to the low frequency of the rs11209026*A allele, only 6 families were informative: two showed transmission (T) and four showed non-transmission (NT) of allele rs11209026*A to affected offspring. Although this observation was in agreement with a protective effect of the rs11209026*A allele, it was not statistically significant due to the small number of cases (data not shown). In celiac disease patients both allele and genotype frequency of the rs11209026 SNP was similar to controls (Table 3).

For the *ATG16L1* rs2241880 SNP we saw an increased frequency of allele G in IBD cases (63%) vs. controls (56.7%) (OR=1.30; CI: 1.10–1.53; p=0.0016) (Table 4);. The association to rs2241880 was mainly due to the CD subgroup (OR=1.36; CI: 1.12–1.66; p=0.0017) and was also significant on the genotype level (p=0.0052) (Table 4). The population attributable risk for carrying allele G in CD is 0.24 and for homozygosity for allele G is 0.19. The association with IBD was confirmed by transmission disequilibrium testing as the rs2241880*G allele was preferentially transmitted to affected individuals (54.7% transmission, OR=1.59; p=0.007) (table 5). We could not find any specific genotype-phenotype associations for *ATG16L1* with respect to disease localization, disease behaviour, need for surgery, occurrence of extraintestinal manifestations and age of onset for both CD and UC.

Table 3. Allele and genotype frequencies of the rs11209026 SNP (*IL23R*) in inflammatory bowel diseases (IBD n=518), Crohn's disease (CD n=311), ulcerative colitis (UC n=207), celiac disease (n=508) and controls (n=893).

	IBD (%)	CD (%)	UC (%)	Celiac (%)	Controls
allele A	10 (1.2)	4 (0.8)	6 (1.9)	76 (7.8)	112 (6.5)
allele G	812 (98.8)	504 (99.2)	308 (98.1)	894 (92.2)	1624 (93.5)
Odds ratio (OR)	0.19	0.14	0.33	1.24	1
95% CI	0.10 - 0.37	0.06 - 0.37	0.15 - 0.73	0.91 - 1.67	–
p-value	6.60E-09	3.90E-07	0.0014	0.17	ref
AA	0 (0)	0 (0)	0 (0)	3 (0.6)	4 (0.5)
AG	10 (2.4)	4 (1.6)	6 (3.8)	70 (14.4)	104 (12.0)
GG	401 (97.6)	250 (98.4)	151 (96.2)	412 (84.9)	760 (87.6)
0	107	57	50	23	25
p-value	5.30E-08	2.40E-06	0.0064	0.4	ref

Table 4. Allele and genotype frequencies of the rs2241880 SNP (*ATG16L1*) in inflammatory bowel diseases (IBD n=518), Crohn's disease (CD n=311), ulcerative colitis (UC n=207), celiac disease (n=508) and controls (n=893).

	IBD (%)	CD (%)	UC (%)	Celiac (%)	Controls
allele G	596 (63.0)	367 (64.2)	229 (61.2)	521 (53.7)	988 (56.7)
allele A	350 (37.0)	205 (35.8)	145 (38.8)	449 (46.3)	754 (43.3)
Odds ratio (OR)	1.30	1.36	1.20	0.89	1
95% CI	1.10 – 1.53	1.12 – 1.66	0.96 – 1.51	0.76 – 1.04	–
p-value	0.0016	0.0017	0.11	0.13	ref
GG	190 (40.2)	121 (42.3)	69 (36.9)	129 (26.6)	280 (32.1)
AG	216 (45.7)	125 (43.7)	91 (48.7)	263 (54.2)	428 (49.1)
AA	67 (14.2)	40 (14.0)	27 (14.4)	93 (19.2)	163 (18.7)
0	45	25	20	23	22
p-value	0.0063	0.0052	0.27	0.09	ref

Table 5. Transmission disequilibrium test results for rs2241880 in ATG16L1 in 230 IBD families (176 complete and 54 incomplete trios; CD=138 and UC=92).

	allele	T	NT	%T	OR	p-value
IBD (CD+UC)	G	180	149	54.7	1.59	0.007
CD	G	109	92	54.2	1.53	0.057
UC	G	71	57	55.4	1.69	0.055

T transmitted allele, NT non-transmitted allele, %T percentage of transmitted alleles. OR odds ratio.

In celiac patients we found no significant association with rs2241880*G,but we did observe a trend for a lower frequency of the rs2241880*G allele in the celiac disease group compared to the controls (53.7% vs. 56.7%) (Table 4).

Discussion

In this study we have confirmed the association of both *IL23R* and *ATG16L1* with CD susceptibility in the Dutch population. We also replicated the association of *IL23R* with the UC phenotype. The two SNPs in *IL23R* and *ATG16L1* did not show association to celiac disease. The rs11209026 SNP in *IL23R* is a non-synonymous SNP (1142G→A) resulting in an amino acid change, Arg381Gln, and the glutamine allele appears to protect against the development of CD. This protective effect has been replicated in an independent cohort very recently and was also confirmed in two independents studies in childhood onset inflammatory bowel diseases. [24-26] The rare A-allele was present in approximately 3% of IBD patients in contrast to 6.0 % of the controls in these studies. In our study we found an even lower frequency of 1.2% in patients compared to 6.5% in controls *IL23R* resides on chromosome 1p31 and encodes a subunit of the proinflammatory cytokine IL23 receptor. IL23 is a heterodimeric cytokine consisting of a p19 and a p40 unit.[27] The finding of *IL23R* as an IBD susceptibility gene is intriguing, since several studies have highlighted the central role of IL23 in gut inflammation. It has been suggested that IL23 initiates and perpetuates both innate- and T-cell-mediated intestinal inflammation.[28] Interestingly, a double-blind controlled study with a human monoclonal antibody against interleukin-12 (IL12) showed a positive effect on induction of remission in patients with active CD. This antibody was directed against the p40 subunit of IL12, which is shared with IL23. Hence, it is tempting to speculate that the observed anti-inflammatory effect was (partly) caused by suppression of IL23 rather then IL12.[29] However, it remains to be elucidated how genetic variants of *IL23R* are involved in IBD susceptibility.

Hampe *et al*. described the association of *ATG16L1* as a susceptibility gene for CD.[16] In our study population we were able to confirm the association of *ATG16L1* with CD. We found a population attributable risk for carrying allele G of 0.24 for IBD and 0.24 for CD, which are similar to the values observed in the original paper.[16] We could not define specific genotype-phenotype associations in our study population for both CD and UC, possibly due to the lack of power. In recent reports the association of *ATG16L1* with CD has been consistently replicated. A large genome wide association study identified the association with the rs2241880 SNP and confirmed this in two independent replication studies.[30] Furthermore, two other studies found similar results.[31 32] In one of these studies a modest association was also found for ulcerative colitis. Although we did not observe a significant association with UC, the trend was the same as for CD, both in the case-control analysis and in the transmission disequilibrium test. Since our UC study group was relatively small (n=207) the association did not reach significant values.

Preliminary data provided by Hampe *et al*. showed expression of the ATG16L1 protein in the intestinal epithelium, but there was no difference in pattern or degree of expression

in tissue from CD patients compared to controls. There was also no difference in protein or cDNA expression in patients carrying different genotypes for rs2241880. ATG16L1 is part of the autophagosome pathway. Autophagocytosis mediates the bulk degradation of cytoplasmic components to the lysosome/vacuole. The process is involved in protein turnover, the starvation response, cellular differentiation and cell death, but also in the processing of intracellular bacteria.[33 34] How this process is implicated in susceptibility for CD remains elusive, although it would support the concept of CD being an inflammatory barrier disorder.[35]

Celiac disease and IBD might share some genetic components, since they share at least two linkage peaks that have been well established for both diseases – on chromosomes 5q31 (CELIAC2, IBD5) and 19p13.1 (CELIAC4, IBD6).[36 37] The discovery that the *MYOB9B* gene is associated with both celiac disease and IBD further suggests that these diseases share part of their genetic susceptibility. We investigated whether *IL23R* and *ATG16L1* variants are common genetic risk factors for both IBD and celiac disease, but found no association of celiac disease with either *IL23R* or *ATG16L1*.

In conclusion, we confirm the genetic association of *IL23R* with both Crohn's disease and ulcerative colitis, and for *ATG16L1* with Crohn's disease, in a Dutch cohort of Caucasian IBD patients. Larger studies are needed to define specific genotype-phenotype associations.

References

1. Podolsky DK. Inflammatory bowel disease. N Engl J Med. 2002 ;347:417-29.

2. Binder VV. Incidence and prevalence of ulcerative colitis and Crohn's disease in the County of Copenhagen, 1962 to 1978. Gastroenterology 1982;83:563-68.

3. Calkins BM, Mendelhoff AI. The epidemiology of idiopathic inflammatory bowel diseases. In Kirsner JB and Shorter RG, eds. Inflammatory Bowel Diseases. 31-68. 1995. Baltimore, Williams & Wilkins.

4. Orholm M, Munkholm P, Langholz E *et al.* Familial occurrence of inflammatory bowel disease. N Engl J Med. 1991 Jan 10;324:84-8

5. Hugot JP, Laurent-Puig P, Gower-Rousseau C *et al.* Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature.1996 29;379:821-3

6. Ogura Y, Bonen DK, Inohara N *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature. 2001;411:603-6.

7. Peltekova VD, Wintle RF, Rubin LA *et al.* Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet. 2004;36:471-5.

8. Silverberg MS, Duerr RH, Brant SR *et al.* Refined genomic localization and ethnic differences observed for the IBD5 association with Crohn's disease. Eur. J. Hum. Genet. 2007;15:328-35.

9. Stoll M, Corneliussen B, Costello CM *et al.* Genetic variation in DLG5 is associated with inflammatory bowel disease. Nat Genet. 2004;36:476-80

10. Schwab M, Schaeffeler E, Marx C *et al.* Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. Gastroenterology. 2003;124:26-33

11. Oostenbrug LE, Drenth JP, de Jong DJ *et al.* Association between Toll-like receptor 4 and inflammatory bowel disease. Inflamm Bowel Dis. 2005;11:567-75.

12. McGovern DP, Hysi P, Ahmad T *et al.* Association between a complex insertion/deletion polymorphism in NOD1 (CARD4) and susceptibility to inflammatory bowel disease. Hum Mol Genet. 2005;14:1245-50.

13. Yamazaki K, McGovern D, Ragoussis J *et al.* Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. Hum Mol Genet. 2005;14:3499-506.

14. Duerr RH, Taylor KD, Brant SR *et al.* A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science. 2006;5804:1461-3.

15. Hue S, Ahern P, Buonocore S *et al.* Interleukin-23 drives innate and T cell-mediated intestinal inflammation. J Exp Med. 2006;203:2473-83.

16. Hampe J, Franke A, Rosenstiel P *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet. 2006 Dec 31; [Epub ahead of print].

17. Weersma RK, Oostenbrug LE, Nolte IM *et al.* Association of interleukin-1 receptor associated kinase M (IRAK-M) and inflammatory bowel diseases. Scand J Gastroenterology 2007; 42; 827-833

18. Report of Working Group of European Society of Paediatric Gastroenterology and Nutrition. Revised criteria for diagnosis of coeliac disease. Arch Dis Child. 1990;65:909-11.

19. van Bodegraven AA, Curley CR, Hunt KA *et al.* Genetic variation in myosin IXB is associated with ulcerative colitis. Gastroenterology. 2006;131:1768-74.

20. International HapMap Consortium. A haplotype map of the human genome. Nature. 2005;437:1299-320.

21. Storm N, Darnhofer-Patel B, van den Boom D *et al.* MALDI-TOF mass spectrometry-based SNP genotyping. Methods Mol Biol.2003;212:241-62.

22. Dudbridge F. Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol. 2003;25:115-21.

23. Haldane JB. The estimation and significance of the logarithm of a ratio of frequencies. Ann Hum Genet. 1956;20:309-11.

24. Tremelling M, Cummings F, Fisher SA *et al*. IL23R variation determines susceptibility but not disease phenotype in inflammatory bowel disease. Gastroenterology. 2007;132:1657-64

25. Dubinsky MC, Wang D, Picornell *et al.*Western Regional Research Alliance for Pediatric IBD. IL-23 receptor (IL-23R) gene protects against pediatric Crohn's disease. Inflamm Bowel Dis. 2007;13:511-5.

26. Van Limbergen JE, Russell RK, Nimmo ER *et al.* IL23R Arg381Gln is associated with childhood onset inflammatory bowel disease in Scotland. Gut. Epub ahead of print, 2 mar 2007: DOI 10.1136/gut2007.122069

27. Oppmann B, Lesley R, Blom B *et al.* Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. Immunity. 2000;13:715-25.

28. Neurath MF. IL-23: a master regulator in Crohn disease. Nat Med. 2007;13:26-8.

29. Mannon PJ, Fuss IJ, Mayer L *et al.* Anti-IL-12 Crohn's Disease Study Group. Anti-interleukin-12 antibody for active Crohn's disease. N Engl J Med. 2004;351:2069-79.

30. Rioux JD, Xavier RJ, Taylor KD *et al.*Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. Nat Genet 2007;39:596-604

31. Prescott NJ, Fisher SA, Franke A *et al* A nonsynonymous SNP in ATG16L1 predisposes to ileal

Crohn's disease and is independent of CARD15 and IBD5. Gastroenterology. 2007;132:1665-71.

32. Cummings JR, Cooney R, Pathan S, *et al.* Confirmation of the role of ATG16l1 as a Crohn's disease susceptibility gene. Inflamm Bowel Dis. 2007 Apr 23; Epub ahead of print DOI 10.1002/ibd.20162

33. Zheng H, Ji C, Li J *et al.*Cloning and analysis of human Apg16L. DNA Seq. 2004;15:303-5.

34. Mizushima N, Kuma A, Kobayashi Y *et al.* Mouse Apg16L, a novel WD-repeat protein, targets to the autophagic isolation membrane with the Apg12-Apg5 conjugate. J Cell Sci. 2003;116:1679-88.

35. Schreiber S, Rosenstiel P, Albrecht M *et al.* Genetics of Crohn disease, an archetypal inflammatory barrier disease. Nat Rev Genet. 2005;6:376-88.

36. Babron MC, Nilsson S, Adamovic S *et al.* Meta and pooled analysis of European coeliac disease data. Eur J Hum Genet. 2003;11:828-34.

37. Rioux JD, Silverberg MS, Daly MJ *et al.* Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. Am J Hum Genet. 2000;66:1863-70.

Chapter 7

An increase in the number of risk-alleles is associated with an increased risk for Crohn's disease and a more severe disease course

R.K. Weersma, P.C.F. Stokkers, A.A van Bodegraven,
R.A. van Hogezand, H.W. Verspaget, D.J.de Jong,
J. van der Woude, B. Oldenburg, R.K. Linskens,
G. van der Steege, D.W. Hommes, J.B. Crusius,
C. Wijmenga, I.M. Nolte, G. Dijkstra
On behalf of the Dutch Initiative on Crohn and Colitis (ICC)

Submitted

Abstract

Background: Inflammatory bowel diseases –consisting of Crohn's disease (CD) and ulcerative colitis (UC)- have a complex genetic background. Several genes have been found to be associated with IBD susceptibility. We performed a large population based case-control study in well phenotyped patients, in order to have enough power to detect these low penetrant genes and to find associations with different subsets of IBD patients. Furthermore, we assessed the risk for developing IBD when combining information from multiple genes.

Methods: 2937 patients with IBD (1696 CD, 1099 UC and 142 Indeterminate Colitis) and 1484 controls from 7 University Medical Centres in the Netherlands were included. A set of 8 SNPs was used for *SLC22A4/5,* 4 SNPs for *DLG5* and 1 SNP for *ATG16L1*. Phenotypic details were available for 2091 patients (1316 CD and 775 UC).

Results: *SLC22A4/5* is associated with IBD but this was not independent from the IBD5 locus. The IBD5 locus was associated with CD with an early age of onset and a complicated disease course. *DLG5* is associated with both CD and UC. *ATG16L1* is associated with CD, specifically with ileal localization and stricturing behavior. Individuals carrying an increasing number of risk alleles were at increasing risk for CD. The progressive increase in Odds Ratios is consistent with an independent effects multiplicative model. CD patients with a more severe disease course or operations had significantly more risk alleles compared to patients without operations or non-stricturing, non-penetrating behavior.

Conclusion: We confirmed the association of the IBD5 locus, *DLG5* and *ATG16L1* with CD and identified specific phenotypes in the Dutch Caucasian population. *DLG5* and the IBD5 locus are also associated with UC indicating that they are a risk factor for IBD in general.
Combining information from the known common risk polymorphisms may enable clinicians to predict the disease course of inflammatory bowel disease patients in the future.

Introduction

Chronic inflammatory bowel diseases (IBD) comprising Crohn's disease (CD) and ulcerative colitis (UC) are characterized by chronic relapsing inflammation of the gastrointestinal tract. The combined prevalence of CD and UC is estimated at 100 – 200 / 100.000 in developed countries.[1 2] Concordance rates in twins and siblings suggest that a genetic predisposition, contributes to the pathogenesis of IBD.[3 4]

In 2001 the *CARD15* gene on chromosome 16 was identified as a susceptibility gene for IBD.[5 6] Three mutations (R702W, G908R, and 1007fsinsC) are independently associated with CD in Caucasian patients. IBD shows a variation in the individual clinical presentation and outcome that is most likely due to differences in the genetic susceptibility in combination with exposure to environmental factors, the commensal bacteria in the intestine, and the intestinal immune system. Since IBD are complex genetic diseases, multiple genes are probably involved in the pathogenesis, each with a small overall contribution to IBD susceptibility. Next to *CARD15* several consecutive genes have been implied to be involved in IBD susceptibility, although none of them has been as consistently replicated as *CARD15*. Three of the most promising genes to be truly associated are *SLC22A4/5* within the IBD5 locus, *DLG5* and *ATG16L1*.[7-9]

A genome wide screen in affected Canadian families identified linkage with a region at chromosome 5q31-q33 (IBD5) with a length of approximately 250 kb conferring susceptibility for CD. [10] Identifying the causative gene(s) in the IBD5 locus has been hampered by the high extent of LD in the region. Nevertheless, two functionally relevant mutations (1672 C→T and -207G→C) in *SLC22A4/5* encoding for the carnitine / organic cation transporter 1 and 2 (OCTN1/2) on the IBD5 locus were associated with CD.[7] Gene-gene interaction with *CARD15* is suggested.[11] Following studies showed contradictory results and the associations found were much weaker than in the original study.[12-14]

By refining their previously defined linkage region on chromosome 10q23, Stoll et al. identified the Drosophila Discs Large Homologue 5 (*DLG5*) gene as a susceptibility gene for CD and IBD.[8] SNP 113 G→A, resulting in an amino acid substitution R30Q, and a low frequent SNP 4136 C→A (P1371Q) were positively associated with IBD and CD. *DLG5* is important in maintaining the epithelial structure and genetic variants could result in impaired intestinal permeability.[15] One study has been able to reproduce these results but other studies from Europe and Japan failed to show any association for DLG5 with IBD.[16-18]

A recent study found a strong association between CD and the autophagy-related 16-like 1 gene (*ATG16L1*) encoding a protein in the autophagosome pathway, which processes intracellular bacteria.[9] A genome-wide association scan in 735 German patients with CD and 368 controls was performed. Analysis of the most informative SNPs in an independent cohort of CD patients and controls identified *ATG16L1* as a CD susceptibility gene. Marker rs2241880, a coding SNP (Thr300Ala), was strongly associated with susceptibility for CD. It was suggested that there is an epistatic effect with one of the three CD associated alleles in *CARD15*.

Many of the replication studies on *SLC22A4/5* and *DLG5* are hampered by the relatively small numbers. To detect low penetrant genes or to have enough power to find associations with different subsets of IBD patients, large population based case-control studies in well pheno-

typed patients are needed. In this study we aimed to answer three important questions. First, we aimed to investigate whether the previously described associations with *SLC22A4/5, DLG5* and *ATG16L1* are present in the Dutch Caucasian population. Second, we studied whether these associations were specific for different phenotypic subsets of IBD patients. Furthermore, we examined the risk for CD when combining information from these three genes. Therefore, we analyzed the association of *SLC22A4/5, DLG5* and *ATG16L1* with IBD, CD, UC and different subsets of CD and UC in a large nationwide case-control study including well phenotyped IBD-patients from seven University Hospitals in the Netherlands.

Methods

Patients and controls

The study population consisted of 2937 patients with IBD (1696 CD, 1099 UC and 142 Indeterminate Colitis) and 1484 controls from seven participating hospitals in the Netherlands: Academic Medical Center, Amsterdam (n=758), VU University Medical Center, Amsterdam (n=636), University Medical Center Groningen, Groningen (n=533), Leiden University Medical Center, Leiden (n=505), Erasmus Medical Center, Rotterdam(n=267), Radboud University Nijmegen Medical Center, Nijmegen (n=157), and University Medical Center Utrecht, Utrecht (n=81).[19-24] Controls consisted of healthy volunteers from the University Medical Center Utrecht (n=868), University Medical Center Leiden (n=168) and from the University Medical Center of Nijmegen (n=139).[19 23 24] The cohort of the University Medical Center Groningen included a large proportion of trios. When DNA of parents was available the non-transmitted haplotypes of each parent were used as a control. When DNA of a child and a spouse was available, both haplotypes of the spouse were regarded as a control (n=309).[21]

Patients were diagnosed on accepted clinical, endoscopic, radiological and histological findings.[1] Almost all individuals studied were of European Caucasian descent. All patients gave informed consent and the study was approved by the institutional ethics review committee of each participating hospital. All DNA samples and data in this study were handled anonymously. For CD patients phenotypic details were registered according to the Vienna classification.[25] For UC patients, phenotypes were described according to age of onset, extend of disease (proctitis, leftsided, or extended), necessity of colectomy and the occurrence of malignancy and extraintestinal manifestations. Phenotypic details were available for 2091 patients (1316 CD and 775 UC) and are presented in table 1 and 2.

Genotyping and SNP Selection

DNA was extracted from 20 ml EDTA-blood following standard procedures and was stored at -80 ^{o}C. Primers to amplify the polymorphic loci were selected using Primer3 software.[26] SNP genotyping was performed as described previously.[21] A set of eight SNPs was analyzed for *SLC22A4/5*. This set included the two known risk-associated SNPs 1672 C→T (rs1050152),

Table 1. Details of patients with Crohn's Disease with phenotypic description available.

Total number	1316		
Sex (m/f)	466 / 850		
Age at diagnosis ¶			
< 40 yrs	1053	(80%)	
> 40 yrs	263	(20%)	
Disease localisation ¶ †			
Ileal	319 / 1316		(24.3%)
Colon	392 / 1316		(29.8%)
Ileocolon	561 / 1316		(42.7%)
Upper GI tract	178 / 1316		(13.5%)
Disease behaviour ¶			
Non-stricturing, non penetraiting	529 / 1307		(40.5%)
Stricturing	345 / 1307		(26.4%)
Penetrating	433 / 1307		(33.1%)
Perianal Disease ¶	275 / 1031		(26.7%)
Extraintestinal Manifestations ¶	244 / 1196		(20.4%)
Operation ¶	604 / 1196		(50.5%)
Family history of IBD ¶	183 / 1316		(13.9%)

¶ Data depicted as cases / total number of patients with specific phenotypic detail available.
† Patients could be classified as having disease localization in the upper GI tract next to ileal, colonic or ileocolonic localization. 44 patients (3.3%) had disease localization in the upper GI tract exclusively.

Table 2. Characteristics of patients with ulcerative colitis with phenotypic description available.

Total number	775	
Sex (m/f)	416 / 359	
Age at diagnosis (years) ¶		
< 40 yrs	545 /769	(70.9%)
> 40 yrs	224 /769	(29.1%)
Disease localisation ¶		
Proctitis	115 / 754	(15.3%)
Leftsided	299 / 754	(39.6%)
Extended	340 / 754	(45.1%)
Extraintestinal Manifestations ¶	95 / 520	(18.3%)
Operation ¶	124 / 619	(20.0%)
Development of malignancy ¶	6 / 619	(1.0%)
Family history of IBD ¶	101 / 775	(13.0%)

¶ Data depicted as cases / total number of patients with specific phenotypic detail available.

Table 3. Marker data for genetic analysis of *SLC22A4/5, DLG5* and *ATG16L1*.

Marker	Type	
SLC22A4/5		
rs7705189	A→G	5' end IBD5
rs3792876	C→T	
rs272893	1082 C→T	
rs1050152	1672 G→C	F503L
rs273900	G→A	
rs2631367	-207 G→C	
rs274551	C→T	
rs2522057	C→G	3' end IBD5
DLG5		
rs2165047	C→T	Haplotype B
rs2289311	A→G	Haplotype A
rs2289310	4136 C→A	Haplotype C; P1371Q
rs1248696	113 G→A	Haplotype D; R30Q
ATG16L1		
rs2241880	G→A	T300A

-207G→C (rs2631367) and SNP (rs3792876) which was found to be strongly associated to pediatric IBD in our institute.[7,26] Additionally three tagging SNPs within the region were analyzed (rs272893, rs273900 and rs274551). Two markers located at the 3' end (rs2522057) and 5' end (rs7705189) were chosen to analyze linkage disequilibrium within the IBD5 region. Four SNPs were analyzed for *DLG5* including the two risk associated SNPs 113 G→A (rs1248696, haplotype D) and 4136 C→A (rs2289310, haplotype C) from the original paper by Stoll *et al*.[8] Two additional haplotype tagging SNPs were chosen to determine haplotype A (rs2289311) and haplotype B (rs2165047). The risk associated SNP rs2241880 was determined for *ATG16L1*.[9] Marker data for all analyzed SNPs are described in table 3.

Statistical methods

As samples from multiple centers were combined, we first tested for homogeneity of the genotype data between centers. Genotype frequencies were compared between the centers using a chi-square test separately for cases and controls. If no homogeneity was observed ($p<0.05$), samples from the center showing the most deviating genotyping frequencies were discarded. Hardy-Weinberg equilibrium was tested among the unrelated unaffected individuals using a chi-square test with 1 degree of freedom. If a marker showed deviation from Hardy-Weinberg equilibrium ($p<0.05$), the marker was discarded from further analyses. The allele and genotype frequencies of patients and controls were compared to test for association using the chi-square test or the Fisher's exact test when appropriate.

Odds ratios (ORs) and 95% confidence intervals (CI) were estimated using binary logistic regression (SPSS 12). Binary logistic regression was also used to determine pairwise and three-way

gene-gene interaction effects with inclusion of the main effects of each gene in the model. In addition, ORs for CD were calculated in a binary logistic regression analysis with number of risk alleles at IBD5 rs2522027, *DLG5* rs2289310 and *ATG16L1* rs2241880 as independent variable compared to controls with zero or one risk allele. Ordinal regression was performed to test whether a more complex course of CD was associated with an increased number of risk alleles. Linkage disequilibrium (LD) parameters D' and r^2 were calculated for all pairs of SNPs within the *SLC22A4/5* gene region and within *DLG5*.
Since all three genes were already identified as candidate genes for IBD, CD and UC, no multiple testing correction is required for testing for susceptibility. Therefore, a multiple testing correction was only performed in the genotype-phenotype subgroup analyses for the number of complementary subgroups of patients using a Bonferroni correction.

Results

Allelic associations for *DLG5, SLC22A4/5* and *ATG16L1* and IBD, CD and UC are summarized in table 4.

SLC22A4/5 / IBD5

The previously described CD associated SNP rs1050152 (1672G→C) was associated with both CD (p=0.046) and UC (p=0.049). Rs2631367 (-207G→C) was not associated with IBD, CD, UC or any subset. Nevertheless, the risk associated TC haplotype for these two SNPs combined was overrepresented in CD patients (41%) and UC patients (41.4%) compared to controls (34.6%). We also observed an association with rs2522057 at the 3'end of IBD5 (OR 1.15; CI 1.05-1.25; p=0.002) with IBD and CD (OR1.15; CI 1.04-1.27; p=0.003) and for rs7705189 at the 5'end of IBD5 with IBD (OR 1.10; CI 0.99-1.23, p=0.040). The association for rs2522057 was strongest for CD with perianal disease (OR 1.37; CI 1.00-1.86; p=0.048) and for carriage of the G-allele for SNP rs7705189 at the 5'end of IBD5 with the necessity of operation in CD (OR 1.32; CI 1.02-1.70; p=0.035). SNP rs273900, which resides between *SLC22A4* and *SLC22A5,* was associated with IBD (OR 1.11; CI 1.00-1.23; p=0.023). An increased risk was found for carriers of the A-allele of rs273900 for a complicated course of CD (penetrating and stricturing disease combined; OR 1.19; CI 1.01-1.41; p=0.039). Carriage of the C-allele for SNP rs3792876 was associated with CD with an early age of onset (OR 12.21; CI 1.27-117.89; p=0.009).
SNPs rs272893, rs273900 and rs274551 showed no homogeneity ($p<0.05$) and the control group of the University Medical Center Groningen was discarded for further analysis for these SNPs. In the remainder of the controls all SNPs satisfied the HWE criteria. All markers were in strong LD (figure 1).

DLG5

Allele C of SNP rs1248696 (R30Q, haplotype D) was associated with IBD (frequencies; OR 1.20; CI 1.05-1.37; p=0.004). This association was present both for CD (OR 1.17; CI 1.01-1.37; p=0.019) and UC (OR 1.25; CI 1.05-1.49; p=0.006). The PAR for carrying allele C was 12.6% for

Table 4. Allelic Association Analysis for *SLC22A4/5*, *DLG5* and *ATG16L1* for Inflammatory Bowel Diseases (IBD, n=2937), Crohn's disease (CD, n=1696) and ulcerative colitis (UC, n=1099).

		Controls	IBD		CD		UC	
SNP ID	**Allele**	**Control/Total (%)**	**Case/Total (%)**	**P value***	**Case/Total (%)**	**P value***	**Case/Total (%)**	**P value***
SLC22A4/5								
rs7705189	G	818/1804 (45,3%)	2463/5160 (47,7%)	**0,040**	1471/3098 (47,5%)	0,074	882/1842 (47,8%)	0,062
rs3792876	C	2404/2622 (91,6%)	5393/5834 (92,4%)	0,115	3068/3340 (91,8%)	0,406	2009/2162 (92,9%)	0,056
rs272893	C	1330/2140 (62,1%)	3702/5734 (64,6%)	**0,024**	2110/3360 (63,9%)	0,091	1380/2118 (65,1%)	**0,021**
rs1050152	C	1663/2888 (57,6%)	3254/5864 (55,5%)	**0,030**	1867/3366 (55,5%)	**0,046**	1199/2170 (55,3%)	**0,049**
rs273900	C	1313/2152 (61,0)	3538/5574 (63,4%)	**0,023**	2013/3186 (63,2%)	0,055	1323/2074 (63,8%)	**0,031**
rs2631367	C	1214/2578 (47,0%)	2304/4710 (48,9%)	0,068	1359/2772 (49,0%)	0,079	844/1730 (48,7%)	0,138
rs274551	C	1846/2204 (83,7%)	4776/5690 (83,9%)	0,423	2755/3254 (84,7%)	0,183	1761/2110 (83,5%)	0,395
rs2522057	C	1197/2904 (41,2%)	2618/5878 (44,5%)	**0,002**	1501/3362 (44,6%)	**0,003**	976/2180 (44,7%)	0,006
DLG5								
rs1248696	G	2501/2880 (86,8%)	5182/5836 (88,7%)	**0,004**	2969/3352 (88,6%)	**0,019**	1923/2156 (89,2%)	**0,006**
rs2289310	C	2479/2618 (94,6%)	5529/5776 (95,7%)	**0,018**	3206/3340 (96,0%)	**0,009**	2001/2102 (95,2%)	0,215
rs2289311	G	675/1054 (64,0%)	1764/2762 (63,9%)	0,460	1047/1628 (64,3%)	0,444	581/930 (62,5%)	0,235
rs2165047	C	2159/2882 (74,9%)	4217/5854 (72,0%)	**0,002**	2401/3352 (71,6%)	**0,002**	1577/2166 (72,8%)	**0,046**
ATG16L1								
rs2241880	G	1515/2664 (56,8%)	3343/5684 (58,8%)	**0,045**	1975/3232 (61,1%)	**0,0005**	1182/2126 (55,6%)	0,185

* one-sided p-value

D' / r^2	rs7705189	rs3792876	rs272893	rs1050152	rs273900	rs2631367	rs274551	rs2522057
rs7705189		0,96	0,96	0,92	0,95	0,95	0,46	0,87
rs3792876	0,07		0,95	0,99	0,94	0,98	0,95	0,87
rs272893	0,46	0,14		0,96	1,00	0,97	0,98	0,92
rs1050152	0,74	0,06	0,41		0,97	1,00	0,97	0,94
rs273900	0,47	0,13	0,96	0,42		0,97	0,98	0,92
rs2631367	0,86	0,08	0,50	0,82	0,52		0,44	0,94
rs274551	0,04	0,02	0,11	0,14	0,11	0,04		0,92
rs2522057	0,64	0,05	0,36	0,88	0,37	0,71	0,13	

Figure 1.
Linkage disequilibrium (LD) across *SLC22A4/5* and the flanking genomic region of the IBD5 locus. D' values for LD are shown for each pair of markers. All markers are in strong LD.

D' / r^2	rs1248696	rs2289310	rs2289311	rs2165047
rs1248696		0,95	0,87	0,97
rs2289310	0,01		0,78	0,97
rs2289311	0,05	0,02		0,84
rs2165047	0,05	0,02	0,14	

Figure 2.
Linkage disequilibrium (LD) across *DLG5*. D' values for LD are shown for each pair of markers. All markers are in LD.

IBD, 6.5% for CD and 22.1% for UC. The association with UC was strongest for patients with leftsided and extensive colitis combined (OR 1.69; CI 1.10-2.61; p=0.014). We did observe an association of SNP rs2289310 (P1371Q, haplotype C) with IBD (OR 1.26;CI 1.01-1.55; p= 0.018) and CD (OR 1.34; CI 1.05-1.71; p=0.009). We found a significant difference in SNP rs2165047 minor allele T frequency (haplotype B) for IBD (OR 0.72; CI 0.56-0.92; p=0.002) and for CD (OR 0.70; CI 0.53-0.91; p=0.002). SNP rs2289311 (haplotype A) was not associated with IBD, CD, UC or any subset.

SNP rs2289310 in the controls from the Radboud University Medical Center, Nijmegen and SNP 2289311 in the cases from the Academic Medical Center, Amsterdam showed deviating genotype frequencies (homogeneity p<0.05) and these specific groups were discarded for further analysis for. All markers were in LD. (figure 2).

ATG16L1

For the *ATG16L1* rs2241880 SNP we observed an increased frequency of allele G in CD cases (61.1%) vs controls (56.8%) (OR=1.19; CI: 1.07–1.32; p=0.0005). The association was also significant on the genotype level (OR=1.46; CI: 1.18-1.81; p=0.0025). The PAR for carrying allele G in CD is 11.6% and for homozygosity for allele G 11.9%. With respect to phenotypic

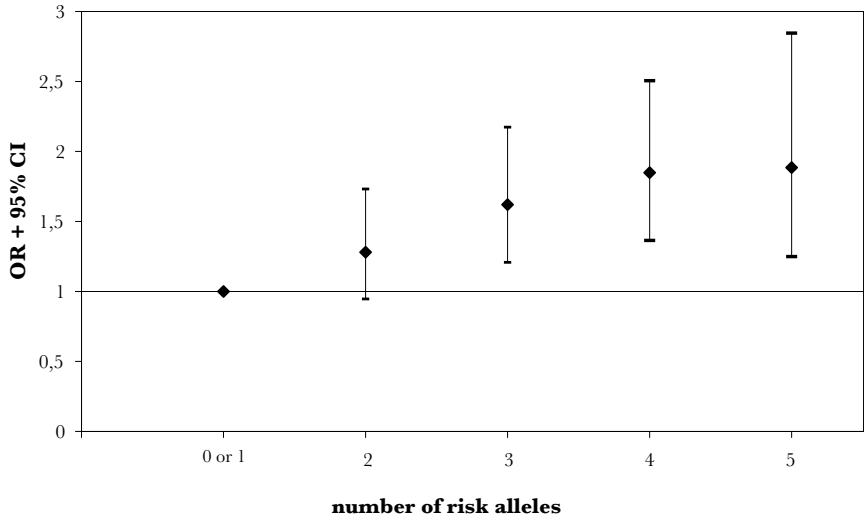


Figure 3.
Odds Ratios (OR) and confidence intervals (CI) for patients carrying increasing numbers of risk alleles for developing Crohn's disease. The risk alleles are IBD5 (the rs2522027*****C allele), *DLG5* (homozygosity for the rs2289310*****C allele) and *ATG16L1* (the rs2241880*****G allele). The reference group consists of individuals with zero or one risk allele. The progressive increase in ORs is consistent with an independent effects multiplicative model.

subgroups we found an allelic association with stricturing behavior of CD (OR=1.27; CI 1.0-1.55; p=0.018) and a tendency to association for carriership of the G-allele with operated CD (OR1.40; CI 1.01-1.95; p=0.067). The rs2241880*G allele was underrepresented in the group with CD with colonic localization and therefore overpresented in the ileocolonic and ileal localized CD groups combined (OR1.20; CI 1.01-1.45; p=0.05)
The rs2241880 showed homogeneous genotype frequencies among the different centers and was in HWE in the combined control group.

Interaction of DLG5, ATG16L1 and IBD5 for Crohn's disease susceptibility
Interaction analysis was performed for the following three CD associated risk genotypes: carriage of the rs2522027*C allele for IBD5, carriage of the rs2241880*G allele for *ATG16L1* and homozygosity for the rs2289310*C allele for DLG5. Pairwise and three-way gene-gene interaction analysis did not show any evidence for statistical interaction on top of the main effects of the individual genes.
Individuals carrying an increasing number of risk alleles were at increasing risk for CD compared to the 9.5% of the controls carrying zero or one risk alleles (Figure 3). This reference

group was chosen because only 0.5% of the controls carried zero risk alleles. In this figure the risk "allele" for DLG5 rs2289310 is the second C allele because the AC genotype was considered the reference genotype due to the low prevalence of the AA genotype. It is shown that with an increasing number of risk alleles the OR for CD susceptibility increases. The progressive increase in ORs is consistent with an independent effects multiplicative model. Next, we used ordinal regression to determine whether an increasing number of risk alleles were associated with a more severe disease course. We found that operated CD patients had significantly more risk alleles compared to patients who were not operated (p=0.0049). Likewise, patients with a more severe disease course (stricturing or penetrating behavior) had more risk alleles compared to patients with non-stricturing, non-penetrating behavior (p=0.0097).

Discussion

It is of utmost importance that previously found genetic associations for IBD are confirmed in independent cohorts with enough power to detect specific genotype-phenotype interactions.[28] To our knowledge, this is the largest replication study that has been performed for the association of *SLC22A4/5, DLG5* and *ATG16L1* with IBD. Furthermore this is the first study to look at the combined impact of three genetic variants associated with CD susceptibility. We showed that the relative risk for CD increases with an increasing number of risk alleles, consistent with an independent effects multiplicative model.

First, we confirmed the association for CD with the IBD5 haplotype. There have been several studies that confirmed the association of the IBD5 haplotype with CD, but identifying the true causal genes has been hampered by the high extend of LD in the region. Peltekova *et al*. found a two locus risk haplotype ("TC haplotype") in *SLC22A4* and *SLC22A5* to be associated independently of the background risk haplotype of IBD5. [7] Although the previously described TC haplotype in *SLC22A4/5* was overrepresented in IBD patients compared to the controls in our study, this effect was not independent of the IBD5 haplotype. For the individual SNPs at these loci we did find an association for rs1050152 (1672G→C) but not for rs2631367 (-207G→C). Our study was sufficiently powered (80% power; p=0.05 in a multiplicative model) to detect an association for these two SNPs in *SLC22A4/5* with an OR of 1.148. Besides this we did find an association for (subsets of) CD with markers at the 5' and 3' end of the IBD5 locus and with a marker in between *SLC22A4* and *SLC22A5.* This indicates that *SLC22A4/5* are associated with CD but that this is not independent from the IBD5 locus. Despite the fact that the true causal gene within the IBD5 locus is not identified yet, many studies have analyzed specific genotype-phenotype associations for IBD5. We observed an association with an early age of onset and a complicated course of CD with specific SNPs in the region. This is in line with several previously published studies reviewed by Reinhard *et al.*[29]

Multiple studies have been performed with conflicting results, regarding the association of the R30Q variant (SNP rs1248696) within *DLG5* with IBD and CD (reviewed by Friedrichs *et al.*)[30] In our study we did find an association for IBD, but surprisingly this risk was higher for UC than for CD, in particular with for UC with an extensive phenotype (e.g. leftsided and

extensive colitis combined). Moreover, this association was observed with the major allele C instead of the minor allele as described by Stoll *et a*l.[8] This is a novel finding as none of the previously published studies have found an association for the R30Q variant with UC.[30] There is substantial heterogeneity for the R30Q variant among different populations, even within Europe, which may be responsible for negative results in previous studies. Furthermore, we did confirm the association of P1371Q (SNP rs2289310) with CD, but could not confirm the protective effect of the haplotype A tagging SNP (rs2289311) as observed in the original paper.[8] Interestingly, we did find a protective effect for both CD and UC for the haplotype B tagging SNP rs2165047, which has not been found to be associated with IBD before. There were no gender-specific effects for *DLG5* in our population as has been observed before.[30] Our results regarding *DLG5* and IBD susceptibility show that there is an association with IBD but this effect is not exclusive for CD. This indicates that *DLG5* is not disease specific but a risk factor for IBD in general.

Furthermore, we could confirm the association of *ATG16L1* with CD. In an initial analysis of the cohort of the University Medical Center Groningen (this thesis, chapter 6), we found a population attributable risk for carrying allele G of 0.24 for CD, which is similar to the value observed in the original paper.[9] When we analyzed the association in the total cohort it still was significantly associated but the PAR fell to 0.116. This is in line with other genetic association studies, where the first study often suggests a stronger effect than is found by subsequent studies.[31] This might be due to bias or due to genuine population diversity. We found the association in particular with CD with ileal or ileocolonic localization, stricturing behavior and a history of surgical intervention. As stricturing behavior is mostly confined to localization in the ileum which frequently results in the necessity of surgery, this is an expected combination. This is the first report of an association of *ATG16L1* with a specific phenotype of CD. We could not find any other associations with UC or any specific subsets,

In this study we show that when combining the information from *ATG16L1*, *DLG5* and the IBD5 region the risk for developing CD increases. There is no specific gene-gene interaction between these three genes but the risk increases with an increasing number of risk alleles. This is the first time that this effect is reported for CD. It is consistent with the idea that multiple genes are involved in CD susceptibility, each with a small overall contribution. We demonstrated that it is possible to combine information from multiple common low-penetrance variants to predict the susceptibility for a complex disease. Moreover, an increasing number of risk alleles was not only associated with CD susceptibility but also with a more complex disease course. Patients with more risk alleles or genotypes had a more severe disease behavior and were more frequently operated. This is an important finding because this emphasizes the fact that it might be possible in the near future to create a genetic risk-profile for a patient to predict the disease course.

In conclusion, we performed the largest case control study for genetic variants in *DLG5, SLC22A4/5* and *ATG16L1* and IBD susceptibility to date. In the Dutch Caucasian population we confirmed the association of *ATG16L1* with Crohn's disease specifically with stricturing behaviour. Secondly, we confirmed the association of *DLG5* with inflammatory bowel diseases, not only with Crohn's disease but also with ulcerative colitis. This indicates that *DLG5* is a risk factor for both phenotypes of inflammatory bowel disease. Thirdly, we found an association of

the IBD5 locus with Crohn's disease which was not independent of the *SLC22A4/5* genes. Most importantly, we showed that when combining the risk-associated variants the relative risk for developing Crohn's disease increases with an increasing number of risk alleles or genotypes consistent with an independent effects multiplicative model. Furthermore, an increase in risk-alleles was associated with a more complex disease course. This finding may enable clinicians to predict the disease course of inflammatory bowel disease patients in the future.

References

1. Podolsky DK. Inflammatory bowel disease. The New England journal of medicine 2002;347:417-29

2. Binder V, V. Incidence and prevalence of ulcerative colitis and Crohn's disease in the County of Copenhagen, 1962 to 1978. Gastroenterology 1982;83:563-68.

3. Orholm M, Munkholm P, Langholz E *et al.* Familial occurrence of inflammatory bowel disease. N Engl J Med. 1991;324:84-8

4. Halfvarson J, Bodin L, Tysk C *et al*. Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. Gastroenterology 2003;124:1767-73.

5. Hugot JP, Laurent-Puig P, Gower-Rousseau C *et al*. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature.1996 29;379:821-3

6. Ogura Y, Bonen DK, Inohara N *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature. 2001;411:603-6.

7. Peltekova VD, Wintle RF, Rubin LA *et al.* Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nature genetics 2004;36:471-75.

8. Stoll M, Corneliussen B, Costello CM *et al.* Genetic variation in DLG5 is associated with inflammatory bowel disease. Nature genetics 2004;36:476-80.

9. Hampe J, Franke A, Rosenstiel P *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. Nat Genet. 2006 Dec 31;

10. Rioux JD, Daly MJ, Silverberg MS *et al.* Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat Genet. 2001;29:223-8.

11. Newman B, Gu X, Wintle R *et al.* A risk haplotype in the Solute Carrier Family 22A4/22A5 gene cluster influences phenotypic expression of Crohn's disease. Gastroenterology. 2005;128:260-9.

12. Torok HP, Glas J, Tonenchi L *et al.* Polymorphisms in the DLG5 and OCTN cation transporter genes in Crohn's disease. Gut 2005;54:1421-7

13. Vermeire S, Pierik M, Hlavaty T *et al.* Association of organic cation transporter risk haplotype with perianal penetrating Crohn's disease but not with susceptibility to IBD.Gastroenterology. 2005;129:1845-53.

14. Noble CL, Nimmo ER, Drummond H *et al.*The contribution of OCTN1/2 variants within the IBD5 locus to disease susceptibility and severity in Crohn's disease. Gastroenterology. 2005;129:1854-64

15. Wakabayashi M, Ito T, Mitsushima M *et al.* Interaction of lp-dlg/KIAA0583, a membrane-associated guanylate kinase family protein, with vinexin and beta-catenin at sites of cell-cell contact J Biol Chem. 2003;278:21709-14.

16. Friedrichs F, Brescianini S, Annese V *et al.*Evidence of transmission ratio distortion of DLG5 R30Q variant in general and implication of an association with Crohn disease in men. Hum Genet. 2006;119:305-1

17. Yamazaki K, Takazoe M, Tanaka T *et al.* Association analysis of SLC22A4, SLC22A5 and DLG5 in Japanese patients with Crohn disease. J Hum Genet 2004;49:664-8.

18. Noble CL, Nimmo ER, Drummond H *et al.* DLG5 variants do not influence susceptibility to inflammatory bowel disease in the Scottish population. Gut. 2005 ;54:1416-20

19. Stokkers PC, Huibregtse K Jr, Leegwater AC, Reitsma PH, Tytgat GN, van Deventer SJ. Analysis of a positional candidate gene for inflammatory bowel disease: NRAMP2. Inflamm Bowel Dis. 2000;6:92-8.

20. van Bodegraven AA, Curley CR, Hunt KA *et al.* Genetic variation in myosin IXB is associated with ulcerative colitis. Gastroenterology. 2006;131:1768-74.

21. Oostenbrug LE, Drenth JP, de Jong DJ *et al.* Association between Toll-like receptor 4 and inflammatory bowel disease. Inflamm Bowel Dis. 2005;11:567-75.

23. Wagtmans MJ, Verspaget HW, Lamers CB, van Hogezand RA. Gender-related differences in the clinical course of Crohn's disease.Am J Gastroenterol. 2001;96:1541-6.

24. de Jong DJ, van der Logt EM, van Schaik A, Roelofs HM, Peters WH, Naber TH. Genetic polymorphisms in biotransformation enzymes in Crohn's disease: association with microsomal epoxide hydrolase.Gut. 2003;52:547-51.

25. Gasche C, Scholmerich J, Brynskov J *et al.* A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. Inflamm Bowel Dis 2000;6:8-15.

26. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365-86.

27. de Ridder L, Weersma RK, Dijkstra G *et al***.** Genetic susceptibility has a more important role in pediatric-onset than in adult-onset Crohn's disease. Inflamm Bowel Dis, 2007 DOI 10.1002/ibd.20171

28. Moonesinghe R, Khoury MJ, Janssens AC. Most published research findings are false-but a little replication goes a long way. PLoS Med. 2007 Feb;4(2):e28.

29. Reinhard C, Rioux JD. Role of the IBD5 susceptibility locus in the inflammatory bowel diseases. Inflamm Bowel Dis. 2006;12:227-38.

30. Friedrichs F, Stoll M. Role of discs large homolog 5. World J Gastroenterol. 2006 ;12:3651-6.

31. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat Genet. 2001;29:306-9.

Chapter 8

Increased incidence of azathioprine-induced pancreatitis in Crohn's disease compared to other diseases

R.K. Weersma, F.T.M. Peters, L.E. Oostenbrug ,
A.P. van den Berg, M van Haastert, R.J. Ploeg, M.D. Posthumus,
J.J. Homan van der Heide, P.L.M. Jansen, H.M. van Dullemen

Alimentary Pharmacology & Therapeutics 2004;20:843-50

Abstract

Background: Azathioprine is widely used in Crohn's disease. A major drawback is the occurrence of side effects, especially acute pancreatitis. Acute pancreatitis is rarely seen when azathioprine is used for other diseases than Crohn's disease.

Methods: The study is a retrospective case note survey of side effects of azathioprine after liver or renal transplantation, for systemic lupus erythemathosis, Wegener's granulomatosis, autoimmune hepatitis, rheumatoid arthritis, ulcerative colitis or Crohn's disease. A computerized search using the term "azathioprine" or "imuran" was done on the Hospital Information System of the university hospital Groningen, resulting in 1564 patients matching our criteria.

Results: Eleven of 224 patients with Crohn's disease experienced acute pancreatitis (4,9%) compared to 2/129 (1,5%) with autoimmune hepatitis, 2/388 (0,5%) after renal transplantation, 1/254 (0,4%) after liver transplantation. Acute pancreatitis was more prevalent in Crohn's disease compared to any other disease. Azathioprine-toxicity necessitating withdrawal occurred significantly ($p<0,05$) more in rheumatoid arthritis (78/317), ulcerative colitis (20/94) and Crohn's disease (52/224) compared to systemic lupus erythemathosis (5/73), Wegener's granulomatosis (6/85), autoimmune hepatitis (8/129), after liver transplantation (17/254) and after renal transplantation (22/388).

Conclusions: Acute pancreatitis is strongly associated with Crohn's disease and rarely occurs with other underlying conditions. Overall azathioprine-induced toxicity and the necessity of withdrawal is more common in inflammatory bowel diseases and rheumatoid arthritis compared to other diseases.

Introduction

Azathioprine is a purine analogue that competitively inhibits the biosynthesis of purine nucleotides. After absorption it is metabolized to 6-mercaptopurine and 6-methyl-mercaptopurine by thiopurinemethyltransferase (TPMT). Azathioprine is widely used in the treatment of inflammatory bowel disease. The occurrence of side effects, however, is a major drawback in the use of azathioprine. Two types of side effects have been reported: the "allergic" non-dose related side effects that include pancreatitis, fever, rash, malaise, hepatitis, nausea and vomiting and the "non-allergic" and dose related side effects as myelosuppression. The genotype of TPMT is an important determinant in the latter type.[1,2]

One of the side effects of azathioprine is acute pancreatitis that usually follows a mild course and occurs in the first 3 to 4 weeks after initiation of treatment. The pathogenesis of azathioprine-induced acute pancreatitis is unknown. It is dose-independent and there is no correlation with myelosuppression, suggesting it is independent of TPMT-levels. A recent population-based case-control study in Denmark found an increased relative risk for pancreatitis in all users of azathioprine.[3] In Crohn's disease this incidence varies between 1,4% and 5% of all patients treated with azathioprine [4-8] and is 3,3% in patients treated with 6-mercaptopurine.[9]

Azathioprine is used in a number of diseases other than inflammatory bowel disease and in transplantation medicine. Surprisingly, acute pancreatitis is rarely seen when azathioprine is used for other diseases than Crohn's disease. In patients with renal transplantation acute pancreatitis is seen in two out of 161 patients treated with azathioprine in one study. [10] Nevertheless, there are a lack of association between the use of azathioprine and acute pancreatitis in renal transplantation patients. Usually, other factors are identified to be responsible for the development of acute pancreatitis.[11,12]

In patients with rheumatoid arthritis, azathioprine is used as a disease-modifying drug. In large studies, no reports have been made of acute pancreatitis as a side effect.[13,14] In patients with autoimmune hepatitis, azathioprine-induced pancreatitis is mentioned very infrequently and the incidence is not exactly known.[15,16] Reviewing studies on systemic lupus erythematosis and lupus nephritis, no cases of pancreatitis were documented in two studies.[17,18] Also, in a recent study with patients with vasculitis associated with antineutrophil cytoplasmic autoantibodies treated with azathioprine, pancreatitis is not mentioned as an adverse event.[19]

In this study we evaluated the occurrence of azathioprine-induced acute pancreatitis in patients with Crohn's disease and compared this to the incidence of acute pancreatitis in other diseases for which azathioprine is indicated. Secondly we evaluated the overall toxicity of azathioprine in Crohn's disease compared to other diseases.

Methods

Patient Selection

The study was performed in a university hospital. It is a retrospective analysis of patient records of patients using azathioprine (Imuran®) for one of the following indications: for rheumatoid arthritis, systemic lupus erythemathosis, Wegener's granulomatosis, autoimmune hepatitis, Crohn's disease, ulcerative colitis, renal transplantation or liver transplantation. Patients used azathioprine on at least two consecutive visits. Patients with an indeterminate diagnosis, more than one diagnosis or with only one visit in the outpatient facility, were excluded. Patients using azathioprine for other indications (e.g.: polymyositis, idiopathic thrombocytopenic purpura, after lung or heart transplantation) were excluded. Data from patients with ulcerative colitis or Crohn's disease from a gastroenterological unit from a community hospital were also separately collected.

Data Collection

Patient data in the University Hospital Groningen were retrieved from the electronic Hospital Information System (HIS). Patient's documents are generated on a word processing system that is used hospital wide. Data are subsequently imported into the DoCma System, which is a permanent data archive, where patient identification is added. A query consisting of the keywords "azathioprine" or "Imuran" was done on the full text of the data stored in the subset internal medicine, using the glimpse tool (http://glimpse.cs. arizona.edu). The query was done on the 29th of August 2002 searching 475.032 documents of 78.906 patients. The specified period was from the 1st of January 1995 till the 31st of July 2002. 2654 patients with charts matching the query were identified. Before inclusion, each patient record was carefully reviewed for the use, or previous use of azathioprine. 67 Patients with Crohn's disease or ulcerative colitis who were using azathioprine from a gastroenterological unit of a community hospital were analyzed separately.

The dosage of azathioprine, occurrence and type of side effects, severe enough to withdraw azathioprine, were noted for all patients. In the case of suspected azathioprine-induced acute pancreatitis, subsequent data were collected. The diagnosis of acute pancreatitis was established by the treating physician at the time. Criteria that were used were elevated amylase in combination of symptoms of upper abdominal or radiating pain, or nausea and vomiting. An asymptomatic rise in amylase levels was not considered an acute pancreatitis. Charts were reviewed on the time of use of azathioprine to onset of symptoms, the need for hospitalization, laboratory data, clinical symptoms of acute pancreatitis, co-medication and other alternative etiologies for acute pancreatitis.

Statistical Analysis

To compare proportions of occurrence of side effects and necessity to withdraw azathioprine in different treatment groups, a chi square test was used with a $p<0,05$ considered as significant.

Table 1. Patient Characteristics - University Hospital.

		RA n=317	SLE n=73	AIH n=129	WG n=85	UC n=94	CD n=224	RT n=388	LT n=254
Dosage	Mean (mg)	116	118	68	90	83	123	83	113
Sex	Male (n)	85	8	27	33	49	79	198	109
	Female (n)	232	65	102	52	45	145	190	145

RA=rheumatoid Arthritis, SLE=Systemic Lupus Erythematosis, AIH=Autoimmune Hepatitis, WG= Wegener's Granulomatosis, UC=Ulcerative Colitis, CD=Crohn's Disease, RT= Renal Transplant LT=Liver Transplant

Results

In the University Hospital the records of 2654 patients matching the query were reviewed. Finally 1564 patients were included. All patients who were treated with Azathioprine in the specified period, for the specified diseases were included. Reasons for exclusion were: treatment for other diseases, patient charts with only one visit to revise or, although azathioprine was mentioned in their charts, patients had not actually been using azathioprine.
Numbers of patients, treated in the university hospital, in each group were: 388 patients after renal transplantation, 317 patients with rheumatoid arthritis, 254 patients after liver patients transplantation, 224 with Crohn's disease, 129 patients with autoimmune hepatitis, 85 patients with Wegener's granulomatosis, 73 patients with systemic lupus erythemathosis, 90 patients with ulcerative colitis and four patients with indeterminate colitis, who were combined with the ulcerative colitis group. Mean dosage and sex distribution are depicted in table 1.

Acute Pancreatitis

In the university hospital acute pancreatitis occurred in eleven of 224 patients with Crohn's disease (4,9%), in two of 129 (1,5%) with autoimmune hepatitis, in two of 388 patients (0,5%) after renal transplantation and in one of 254 patients after liver transplantation (0,4%). Azathioprine-induced pancreatitis was not seen in patients with systemic lupus erythemathosis, Wegener's granulomatosis or rheumatoid arthritis. Acute pancreatitis occurred more in patients with Crohn's disease than in any other treatment group ($p<0,05$). Only in the comparison with autoimmune hepatitis the difference did not reach a statistical significance ($p=0.11$).
In the community hospital azathioprine-induced pancreatitis occurred in 2/41 patients (4.9%) with Crohn's disease and in 1/26 patients (3.8%) with ulcerative colitis.
Characteristics of patients with azathioprine-induced pancreatitis in the university hospital and the community hospital combined, are shown in table 2. In all patients the treating physician diagnosed azathioprine-induced pancreatitis according to the rapid onset of symptoms and at least a 3-fold rise in serum amylase after initialising azathioprine. Symptoms resolved in all patients after withdrawal of azathioprine.

Table 2. Side-effects of Azathioprine necessitating withdrawal - University Hospital.

	RA n=317	**SLE** n=73	**AIH** n=129	**WG** n=85	**UC** n=94	**CD** n=224	**RT** n=388	**LT** n=254	**Total** n=1564	**Total** (%)
Side-effects (n)										
nausea and vomiting	33	2	2	2	5	20			64	4,1
abdominal pain	8				4	6			18	1,2
diarrhea	3		1			1			5	0,3
fever	8		2	4	6	5		1	26	1,7
rash	2	1	2			3			8	0,5
hepatitis	22	1	1		2	5	6	6	43	2,7
myelosuppression	11	3	1		1	4	10	8	38	2,4
arthralgia	3		1			8			12	0,8
general weakness	5			3		3			11	0,7
acute pancreatitis			2			11	2	1	16	1,0
infections	3		1			1			5	0,3
other	3		1		1	3	5		13	0,8
withdrawn(n)	78	5	8	6	20	52	22	17	208	
withdrawn(%)	24,6	6,8	6,2	7,1	21,3	23,2	5,7	6,7	13,2	

Denotates all side-effects necessitating withdrawal. One patient can have more then one side-effect. RA=rheumatoid Arthritis, SLE=Systemic Lupus Erythematosis, AIH=Autoimmune Hepatitis, WG= Wegener's Granulomatosis, UC=Ulcerative Colitis, CD=Crohn's Disease, RT= Renal Transplant, LT=Liver Transplant

Table 3. Side-effects of Azathioprine in IBD; university vs community hospital.

	University Hospital		Community Hospital		Total	
	CD	**UC**	**CD**	**UC**	**CD**	**UC**
	n=224	n=94	n=41	n=26	n=265	n=120
Side-effects (n)						
nausea and vomiting	20	5	2	2	22	7
abdominal pain	6	4	2		8	4
diarrhea	1				1	
fever	5	6	2		7	6
rash	3				3	
hepatitis	5	2	3	1	8	3
myelosuppression	4	1			4	1
arthralgia	8				8	
general weakness	3				3	
acute pancreatitis	11		2	1	13	1
infections	1				1	
other*	3	1			3	1
withdrawn(n)	52	20	9	4	61	24
withdrawn(%)	23,2	21,3	22,0	15,4	23,0	20,0

Denotates all side-effects necessitating withdrawal. One patient can have more then one side-effect. CD=Crohn's Disease; UC=Ulcerative Colitis

Mean amylase level in all 13 patients with Crohn's disease was 1505 U/l and mean CRP level was 75. Lipase levels are not routinely measured in both hospitals. Mean dosage of azathioprine was 123 mg (1.75 mg / kg). Only one of 13 patients used alcohol in an amount of 2 units a day. In 8 of 13 patients a biliary origin of the pancreatitis was actively excluded by ultrasonography. Since in all 13 patients symptoms resolved soon after withdrawal of azathioprine no further abdominal CT scans were deemed necessary by the treating physician at time of diagnosis.

None of the 13 patients experienced an episode of pancreatitis before the treatment of azathioprine. None of the 13 patients used sulfasalazin. Six of 13 patients were using mesalazine at the moment azathioprine was started. All six were using mesalazine for several years without symptoms of acute pancreatitis and continued to do so after withdrawal of azathioprine.

For all 19 patients who developed pancreatitis, the mean time to onset of symptoms was 21 days, excluding one patient after liver transplantation. In all patients symptoms resolved after withdrawing azathioprine. Fifteen of nineteen patients required hospitalization. No patient experienced a rash or raised eosinophil count. None of the patients developed leucopenia

Only two patients were rechallenged with azathioprine. One patient with Crohn's disease developed symptoms of acute pancreatitis after two days and one patient with renal transplant and presumed azathioprine-induced pancreatitis developed no symptoms and tolerated azathioprine after rechallenge.

Side effects other than acute pancreatitis

Frequency and type of side effects, severe enough to withdraw azathioprine, are shown in table 3. One patient can have more than one type of side effect. The most common side effects of all patients combined were nausea and vomiting (4,1%), hepatitis (2,7%), myelosuppression (2,4 %), fever (1,7%), abdominal pain (1,2%) and acute pancreatitis (1,0%).

Side effects depicted as "other" included: anxiety (one patient with rheumatoid arthritis), verrucae (one patient with rheumatoid arthritis) macrocytosis (one after liver transplantation, one after renal transplantation), pulmonary embolism (one patient with ulcerative colitis), neuro-toxicity (two patients with Crohn's disease), veno-occlusive disease (one patient after renal transplantation), weight loss (one patient with systemic lupus erythemathosis) and not specified (one rheumatoid arthritis, one Crohn's disease, one autoimmune hepatitis and one patient after renal transplantation).

Withdrawal due to side effects occurred more often ($p<0,05$) in patients with inflammatory bowel disease (Crohn's disease 23,2% and ulcerative colitis 21,3%) and rheumatoid arthritis (24,6%) than in patients with systemic lupus erythemathosis (6,8%), autoimmune hepatitis (6,2%), Wegener's granulomatosis (7,1%) after liver transplantation (6,7%) or renal transplantation (5,7%).

Nausea and vomiting as a reason to discontinue azathioprine occurred more often in rheumatoid arthritis (10,4%), Crohn's disease (8,9%) and ulcerative colitis (5,3%) compared to systemic lupus erythemathosis (2,7%), Wegener's granulomatosis (2,4%), autoimmune hepatitis (1,6) or after renal transplantation (0%) and liver transplantation (0%). Hepatitis during the use azathioprine, necessitating withdrawal occurred more often ($p<0,05$) in rheumatoid arthritis (6.9%) than in Crohn's disease (2,2%), ulcerative colitis (2,1%), systemic lupus erythemathosis (1,4%), autoimmune hepatitis (0,8%) after renal transplantation (1,4%) or after liver transplantation (2,4%). Myelosuppression as the reason for withdrawal occurred more often in rheumatoid arthritis (3,5%) systemic lupus erythemathosis (4,1%), after liver transplantation (3,1%) or after renal transplantation (2,6%), compared to Crohn's disease (1,8%), ulcerative colitis (1,0%), autoimmune hepatitis (0,8%) and Wegener's granulomatosis (0%).

In patients treated for inflammatory bowel diseases, there were no differences in the percentage of withdrawal between the University Hospital and the community hospital. In the University Hospital 52/224 (23.2%) patients treated for Crohn's disease had to stop because of side effects versus 9/41 (22.0%) patients in the community hospital. For ulcerative colitis the numbers were 20/94 (21.3%) and 4/26 (15.4%) (Table 4).

Table 4. Characteristics of patients with azathioprine-induced acute pancreatitis university and community hospital combined.

	CD n=13	UC n=1	AIH n=2	RT n=2	LT n=1
patient					
mean age (yrs)	39	64	62	44	40
sex	10 f / 3 m	1 f	2 f	1 f / 1 m	1 f
mean dosage (mg/day)	123	125	75	88	125
time to onset (days)	25	14	14	9	2550
hospitalisation	11/13	1/1	2/2	1/2	-*
rechallenge (n)	1/13	0	0	1/2	0
Recurrence of symptoms	yes	-	-	no	-
symptoms:					
upper abdominal pain	11/13	1/1	2/2	1/2‡‡	1/1
nausea	10/13	1/1	0/1	1/2‡‡‡	1/1
vomiting	7/13	1/1	0	0/2‡‡‡	1/1
radiating pain	4/13	1/1	2/2	1/2‡‡‡	0
fever	4/13	1/1	0	1/2‡‡‡	1/1
laboratory data					
Mean CRP (mg/l)	75	181	10	136	-*
Mean serum amylase (U/l)	1504	70 †	1937	1857	-*
Localisation (CD)					
colon	13/13	-	-	-	-
term ileum	9/13	-	-	-	-
other	0/13	-	-	-	-

* data not available; † urine amylase 1350 U/l ; ‡ ‡ ‡data not avalailable for second patient.
CD=Crohn's Disease, UC=Ulcerative Colitis, AIH=Autoimmune Hepatitis, RT=Renal Transplant, LT=Liver Transplant,

Discussion

In this study 1564 patients treated with azathioprine were reviewed for the occurrence of acute pancreatitis and other side effects necessitating withdrawal. The retrospective nature of the data makes interpretation difficult, but to our knowledge this is the only comparison of the use of azathioprine in different diseases and secondly, it reflects the use of azathioprine in clinical practice.

In accordance with data from the literature, acute pancreatitis occurred significantly more often in patients with Crohn's disease than in all other groups, suggesting a relation with the underlying condition. Frick et al. described a lack of association between azathioprine and

pancreatitis in renal transplantation medicine.[12] Also in our group of patients after renal transplantation, the presumed association was not evident. The first patient did not develop symptoms after rechallenge with azathioprine. The second patient experienced acute pancreatitis attributed to the use of tacrolimus, which was withdrawn. Symptoms of severe acute pancreatitis returned when azathioprine was started consecutively. These findings again suggest that azathioprine is probably only a co-factor in inducing pancreatitis in renal transplantation patients. Furthermore, symptoms of acute pancreatitis became apparent only after 7 years in 1 patient with liver transplantation, which makes the relation with azathioprine highly unlikely in this patient.

Average duration of treatment in patients with Crohn's disease and azathioprine-induced pancreatitis was 25 days, meaning a strong temporal relationship between the start of azathioprine treatment and the emergence of symptoms of pancreatitis. Although patients in a University Hospital are highly selected, a selection bias because of referral doesn't seem to be a good explanation of the high incidence of azathioprine induced pancreatitis in Crohn's disease. All but one patients with pancreatitis started treatment with azathioprine in the university hospital. Secondly the incidence is the same as in earlier studies. And finally, selection bias does not explain the difference in incidence of pancreatitis in comparison with the other disease populations, who were also treated in the University Hospital.

Symptoms of acute pancreatitis required hospitalization in 78% of all cases, which was never necessary for the large group of patients with abdominal pain, nausea and vomiting due to azathioprine in the other treatment groups. Thus selection bias by gastroenterologists, who might consider acute pancreatitis earlier than other medical specialists, is unlikely.

The use of aminosalicylates is also associated with acute pancreatitis and the simultaneous use of aminosalicylates and azathioprine is associated with an inhibition of TPMT activity resulting in an increased production of 6-thioguanine nucleotide, that could result in a higher risk of bone marrow suppression.[20-22] Six patients were using mesalazine at the moment azathioprine was started. The simultaneous use of aminosalicylates cannot be ruled out to be a contributing factor to the occurrence of acute pancreatitis, although all six were using mesalazine for several years without symptoms of acute pancreatitis and continued to do so after withdrawal of azathioprine. None of the 6 patients using both drugs developed leucopenia.

The pathogenesis of azathioprine-induced acute pancreatitis is unknown. It is dose-independent and there is no correlation with myelosuppression, suggesting it is independent of TPMT-levels. A delayed type II or IV allergic reaction has been suggested.[5] The timing of symptoms is compatible with the development of antibodies. The recurrence of symptoms within several hours after rechallenge supports this idea. However our analysis suggests that there is a correlation with Crohn's disease, which suggests a disease specific underlying mechanism. An immune mediated idiosyncratic drug reaction due to a genetic predisposition might be an explanation.

Circulating pancreatic antibodies (PAB) are found in approximately 30 % of patients with CD (table1). [23-25] PAB are not found in healthy controls, or in patients with other gastrointestinal diseases and in various autoimmune disorders (including autoimmune hepatitis, systemic

lupus erythematosis and rheumatoid arthritis). Since PAB and azathioprine-induced pancreatitis are both specific for Crohn's disease, an association or pathogenic role of PAB in azathioprine-induced pancreatitis can be hypothesized. One study suggests that patients with CD and pancreatic exocrine insufficiency were significantly more likely to be PAB positive than patients with CD without pancreatic insufficiency.[26] The reason why CD patients have pancreatic insufficiency is unknown, but it might be due to a low-grade inflammation of the pancreas, possibly as an extraintestinal manifestation of CD. Furthermore, the use of azathioprine worsens the inflammation of the pancreas which was shown in an animal model of acute pancreatitis in rats.[27] Extrapolating these findings one could suggest that adding azathioprine to an already inflamed pancreas in PAB positive CD patients aggravates the inflammation and leads to a clinical overt picture of acute pancreatitis. In addition to our retrospective case note survey we obtained eight samples of CD patients with azathioprine-induced pancreatitis and compared the occurrence of circulating PAB (determined by a standardized immunofluorescence assay) with 26 CD patients not using azathioprine.[28] 25% of identified patients with azathioprine induced pancreatitis had circulating PAB versus 8% of controls with CD. Although there was a difference, this was not statistically significant. PAB titers in the controls were very low, whereas PAB were detectable in high concentrations in the patients with pancreatitis. This may indicate that the presence of PAB at high titers increases the risk but is not a determining factor for developing pancreatitis.

Besides the occurrence of acute pancreatitis, the overall azathioprine-induced toxicity varies between different diseases. In our study withdrawal due to side effects was significantly higher in patients with rheumatoid arthritis, Crohn's disease and ulcerative colitis compared to systemic lupus erythemathosis, Wegener's granulomatosis, autoimmune hepatitis, renal transplantation and liver transplantation.

Difference in interpretation of symptoms and the experience with the use of azathioprine by different medical specialists may vary and might be an explanation for the disparity in withdrawal during the treatment of different diseases. However, there is frequent communication between specialists in gastroenterology, hepatology, and transplantation medicine. Similarly, there is close contact between specialists in rheumatology, immunology and nephrology. Furthermore, the percentage of withdrawal in patients with Crohn's disease and ulcerative colitis did not differ between the university and the community hospital. Therefore, differences in experience and interpretation of symptoms do not seem to be a satisfying explanation for the difference in appreciated adverse events.

Since our aim was to detect a difference in occurrence of azathioprine-induced acute pancreatitis, we did not collect data on medication used simultaneously at the time of occurrence of other side effects than pancreatitis. The concurrent use of steroids might be a contributing factor for the differences. However, there is a difference in percentage of withdrawal in systemic lupus erythemathosis, autoimmune hepatitis, ulcerative colitis, Crohn's disease and Wegener's granulomatosis, while most of these patients use steroids when azathioprine is started.

In conclusion, there is a clear difference in percentage and severity of azathioprine-induced toxicity in different diseases. The occurrence of acute pancreatitis due to treatment with azathioprine is strongly associated with Crohn's disease and rarely occurs in other diseases.

The reason for this is unknown but it is not associated with the occurrence of circulating pancreatic antibodies. An idiosyncratic drug reaction due to a genetic predisposition is supposed. The necessity to withdraw azathioprine is more frequent in inflammatory bowel diseases and rheumatoid arthritis compared to Wegener's granulomatosis, systemic lupus erythemathosis, autoimmune hepatitis and after renal transplantation or liver transplantation.

References

1. Krynetski EY, Tai HL, Yates CR *et al.* Genetic polymorphism of thiopurine S-methyltransferase: clinical importance and molecular mechanisms. Pharmacogenetics 1996;6:279-90.

2. Yates CR, Krynetski EY, Loennechen T *et al.* Molecular diagnosis of thiopurine S-methyltransferase deficiency: genetic basis for azathioprine and mercaptopurine intolerance. Ann Intern Med 1997;126:608-14.

3. Floyd A, Pedersen L, Nielsen GL, Thorlacius-Ussing O, Sorensen HT. Risk of acute pancreatitis in users of azathioprine: a population-based case-control study. Am J Gastroenterol. 2003;98:1305-8.

4. Summers RW, Switz DM, Sessions JT Jr *et al.* National Cooperative Crohn's Disease Study: results of drug treatment. Gastroenterology 1979;77:847-6.

5. Sturdevant RA, Singleton JW, Deren JL, Law DH, McCleery JL. Azathioprine-related pancreatitis in patients with Crohn's disease. Gastroenterology 1979;77: 883-886.

6. Sandborn W, Sutherland L, Pearson D, May G, Modigliani R, Prantera C. Azathioprine or 6-mercaptopurine for inducing remission of Crohn's disease. Cochrane Database Syst Rev 2000;(2): CD000545.

7. Present DH, Meltzer SJ, Krumholz MP, Wolke A, Korelitz BI. 6-Mercaptopurine in the management of inflammatory bowel disease: short- and long-term toxicity. Ann Intern Med 1989;111:641-95.

8. Kirschner BS. Safety of azathioprine and 6-mercaptopurine in pediatric patients with inflammatory bowel disease. Gastroenterology 1998;115:813-21.

9. Haber CJ, Meltzer SJ, Present DH, Korelitz BI. Nature and course of pancreatitis caused by 6-mercaptopurine in the treatment of inflammatory bowel disease. Gastroenterology 1986;91:982-6.

10. Kahan BD. Efficacy of sirolimus compared with azathioprine for reduction of acute renal allograft rejection: a randomised multicenter study. The Rapamune US Study Group. Lancet. 2000;356:194-202.

11. Frick TW, Fryd DS, Sutherland DE, Goodale RL, Simmons RL, Najarian JS. Hypercalcemia associated with pancreatitis and hyperamylasemia in renal transplant recipients. Data from the Minnesota randomized trial of cyclosporine versus antilymphoblast azathioprine. Am J Surg 1987;154:487-9.

12. Frick TW, Fryd DS, Goodale RL, Simmons RL, Sutherland DE, Najarian JS. Lack of association between azathioprine and acute pancreatitis in renal transplantation patients. Lancet 1991;337:251-2.

13. Singh G, Fries JF, Spitz P, Williams CA.Toxic effects of azathioprine in rheumatoid arthritis. A national post-marketing perspective. Arthritis Rheum 1989;32:837-43.

14. Suarez-Almazor ME, Spooner C, Belseck E. Azathioprine for treating rheumatoid arthritis. Cochrane Database Syst Rev 2000;(4):CD001461.

15. Kanzler S. Long-term prognosis and side effects of immunosuppressive therapy in 122 patients with autoimmune hepatitis (AASLD abstract). Hepatology 1997;26:535A.

16. Johnson PJ, McFarlane IG, Williams R. Azathioprine for long-term maintenance of remission in autoimmune hepatitis. N Engl J Med 1995;333:958-63.

17. Chan TM, Li FK, Wong RW, Wong KL, Chan KW, Cheng IK. Sequential therapy for diffuse proliferative and membranous lupus nephritis: cyclophosphamide and prednisolone followed by azathioprine and prednisolone. Nephron. 1995;71:321-7.

18. Chan TM, Li FK, Tang CS *et al.* Efficacy of mycophenolate mofetil in patients with diffuse proliferative lupus nephritis. Hong Kong-Guangzhou Nephrology Study Group. N Engl J Med. 2000;343: 1156-62.

19. Jayne D, Rasmussen N, Andrassy K *et al.* A randomized trial of maintenance therapy for vasculitis associated with antineutrophil cytoplasmic autoantibodies. N Engl J Med. 2003 ;349:36-44.

20. Deprez P, Descamps C, Fiasse R. Pancreatitis induced by 5-aminosalicylic acid. Lancet. 1989;2:445-6.

21. Szumlanski CL, Weinshilboum RM. Sulphasalazine inhibition of thiopurine methyltransferase: possible mechanism for interaction with 6-mercaptopurine and azathioprine. Br J Clin Pharmacol. 1995;39:456-9.

22. Dewit O, Vanheuverzwyn R, Desager JP, Horsmans Y. Interaction between azathioprine and aminosalicylates: an in vivo study in patients with Crohn's disease. Aliment Pharmacol Ther. 2002;16:79-85.

23 Seibold F, Weber P, Jenss H, Wiedmann KH. Antibodies to a trypsin sensitive pancreatic antigen in chronic inflammatory bowel disease: specific markers for a subgroup of patients with Crohn's disease. Gut. 1991 ;32:1192-7

24 Joossens S; Vermeire S; Van Steen S *et al*. Pancreatic Autoantibodies in Inflammatory Bowel Disease. Inflammatory Bowel Diseases. 2004.10:771-777

25 Klebl FH, Bataille F, Huy C, Hofstadter F, Scholmerich J, Rogler G. Association of antibodies to exocrine pancreas with subtypes of Crohn's disease. Eur J Gastroenterol Hepatol. 2005;17:73-7

26 Seibold F, Scheurlen M, Muller A, Jenss H, Weber P. Impaired pancreatic function in patients with Crohn's disease with and without pancreatic autoantibodies. J Clin Gastroenterol. 1996;22:202-6.

27 Foitzik T, Forgacs B, Ryschich E *et al*. Effect of different immunosuppressive agents on acute pancreatitis: a comparative study in an improved animal model.Transplantation 1998;65:1030-6.

28 Winnock F, Christie MR, Batstra MR *et al.* Belgian Diabetes Registry. Autoantibodies to a 38-kDa glycosylated islet cell membrane-associated antigen in (pre)type 1 diabetes: association with IA-2 and islet cell autoantibodies. Diabetes Care. 2001;24:1181-6.

Chapter 9

Summary and future perspectives

Summary

Chronic inflammatory bowel diseases (IBD) comprising Crohn's disease (CD) and ulcerative colitis (UC) are characterized by chronic relapsing inflammation of the gastrointestinal tract. The combined prevalence of CD and UC is estimated at 100 – 200 / 100.000 in developed countries. The pathogenesis of IBD is only partially understood but concordance rates in twins and siblings suggest that a genetic predisposition, apart from environmental and immunological factors, contributes to the pathogenesis of IBD. In the past decade, tremendous progress has been achieved in unraveling the genetic etiology of IBD; several susceptibility loci for IBD have been identified and the *CARD15* gene encoding for the NOD2 protein on chromosome 16 (IBD1) has been found to be strongly related to CD susceptibility. NOD2 is part of the innate immune system and recognizes specific bacterial membrane components. Two missense mutations (R702W and G908R) and one frameshift insertion mutation (L1007fsinsC) in the leucine rich repeat region of the protein are independently associated with ileal CD in caucasian patients, suggesting that a defect in bacterial recognition might be involved in CD. Since CD and UC have many different phenotypic presentations and genome wide scans detected several linkage regions, IBD is considered a multigenic disease . Multiple low penetrance genes, each with a small overall contribution are probably associated with (different subsets of) UC and CD.

For genetic research in complex genetic disorders as IBD it is important to have large homogenous cohorts of well described patients. In 2001 the CODE study (Chronische Ontsteking van de Darm en Erfelijkheid: Chronic Inflammation of the Gut and Inheritance) started in the University Medical Center Groningen. DNA has been, and is still being collected of a large cohort of mainly Caucasian patients and family members from the Northern part of the Netherlands. An important aspect in studying IBD genetics is the consequent description of disease phenotypes. Since IBD is considered a multigenic disorder, different genes are probably involved in different subsets of phenotypes. It is therefore mandatory to have internationally accepted classification systems for IBD. For the current thesis, the Vienna classification is used, an internationally accepted and frequently used system.

Another research subject in IBD genetics is pharmacogenetics. There has been much interest in the pharmacogenetics of azathioprine metabolism. Azathioprine is a purine analogue that is frequently used in the treatment of Crohn's disease but its use is hampered by the frequent occurrence of side-effects. Polymorphisms in the thiopurinemethyltransferase (TPMT) gene, which metabolizes azathioprine to 6-mercaptopurine and 6-methyl-mercaptopurine, and inosine triphosphate pyrophospatase (ITP-ase) deficiency which leads to accumulation of the metabolite 6-thio-ITP, have been found to be responsible for a subset of the side-effects of azathioprine therapy.

This thesis focuses on the genetic background of IBD. Our aim was to find specific genotype-phenotype associations for previously identified genes and to find novel IBD associated genes. The first part explores the association of two novel genes with IBD. The middle part of the thesis consists of three studies replicating previously described associations in specific cohorts using different approaches and the last part focuses on azathioprine toxicity and specific diseases or phenotypes.

Chapter 1 starts with an introduction and describes the aims and outlines of the thesis.

This thesis focus on specific genetic associations with IBD. Therefore, a detailed review of the current literature on IBD genetics is given in **Chapter 2.** Tremendous progress has been achieved in unraveling the genetic etiology of IBD. It has led to the discovery of mutations in *CARD15* encoding for NOD2, associated with ileal CD. However it is only partially understood how mutations in *CARD15* lead to CD. Mouse models, in vitro data and studies in humans offer conflicting data whether there is a loss or gain of function of NOD2 in CD. Current research on the role of mutations in *CARD15* in the pathogenesis of IBD is reviewed. Since the discovery of the association of *CARD15* and CD, many additional genes have been studied. Several of these genes are potentially truly associated, but results have been conflicting for many of the associations found. Promising candidate genes include Toll like receptor 4 (*TLR4*), Multi Drug Resistance 1 (*MDR1*), NOD1 (*CARD4*), *HLA DRB1*103*, *DLG5* as well as the IBD5 locus on chromosome 5, including members of the organic cation transporter cluster 1 and 2. The current literature for IBD associated genes is extensively reviewed. Furthermore, once a genetic association is established, it is important to identify specific phenotypes or subgroups that are associated with specific genetic variants. Therefore a uniform classification scheme like the Vienna classification (or its recent modification the Montreal Classification) is of utmost importance.

The first part of the thesis comprises two chapters studying the possible association between novel candidate genes and IBD susceptibility. We performed a case control study for the association of the gene encoding for the interleukin-receptor associated kinase-M (*IRAK-M*) and IBD susceptibility in **chapter 3.** IRAK-M is a Nuclear Factor κB (NF-κB) dependent, negative regulator of Toll-like receptor signaling. Toll-like receptors (TLRs) are, like NOD2, important components of the innate immune system. TLRs have extracellular leucine-rich-repeat domains that recognize pathogen associated molecular patterns (PAMPs) and activate adaptive immunity. Chronic repeated stimulation of TLRs by PAMPs causes an "endotoxin tolerance". This process involves different mechanisms, including TLR4 downregulation and decreased NF- κB activation. In a previous study it was shown that genetically engineered *IRAK-M* knockout mice have an impaired endotoxin tolerance and show an increased inflammatory responses to bacterial infection. The function of IRAK-M as a negative regulator of TLR signaling, inducing an impaired endotoxin tolerance and an inflammatory response, in combination with its localization at an IBD susceptibility locus (IBD2) makes *IRAK-M* a good candidate gene for susceptibility for inflammatory bowel diseases. Therefore, we studied the association of *IRAK-M* with IBD in a cohort of 542 patients with IBD (309 CD, 233 UC) and 305 controls. Two exonic single nucleotide polymorphisms (SNPs) and six microsatellite markers were evaluated by allelic association analysis and the haplotype sharing statistics. Results were stratified for the three known *CARD15* mutations R702W, G908R and 1007fsinsC. SNP genotyping was carried out by using TaqMan PCR primer/probe sets, designed through Applied Biosystems' Assay by Design service by a standardized protocol. Similarly, a standardized procedure was used for the microsatellite marker polymerase chain reactions.
Power analysis showed that an association with an odds ratio (OR) of >1.75 should be de-

tected in our study population. However, we observed no significant differences in the distribution of *IRAK-M* allele frequencies between the patients with IBD, CD or UC and controls. Analysis of specific phenotypes of CD and UC did not reveal any significant differences either. Carriership of one or more of the three IBD associated mutations in *CARD15* in combination with a specific microsattelite marker showed an increased risk for UC compared to patients not carrying a *CARD15* mutation. Nevertheless, since *CARD15* has not been shown to be involved in UC susceptibility before, it is doubtful whether this is a true association or merely a false positive result.

In **Chapter 4** the genetic association between *RUNX3* and IBD is studied. RUNX3 is a member of the runt domain family of transcription factors which are increasingly being recognized to be involved in autoimmunity. Several reasons make *RUNX3* a good candidate gene for IBD susceptibility. Firstly, loss of Runx3 function is associated with a spontaneous colitis in knockout mice. Secondly, RUNX3 is a member of the TGF- ß signaling pathway, which is a potent inhibitor of inflammation in IBD. Impaired activation of RUNX3 might result in decreased activity of the TGF- ß pathway and decreased inhibition of inflammation in IBD. Finally the gene encoding for *RUNX3* resides on chromosome 1p36, which is a susceptibility locus for IBD. Four SNPs and four microsattelite markers were studied for *RUNX3* in the CODE cohort. Furthermore, mutations in the genes *SLC22A4* and *5* encoding for the organic cation transporters 1 and 2 (OCTN1/2) were found to be associated with CD in a previous publication. Another publication showed an association between polymorphisms in *SLC22A4*, resulting in a disrupted binding site for RUNX and rheumatoid arthritis. For that reason, six SNPs in *SLC22A4/5* (including the known polymorphisms -207 G→C, 1672 C→G and the SNP disrupting the RUNX binding site in *SLC22A4*) were analyzed for association with IBD and interaction with *RUNX3* was studied. All results were stratified for *CARD15* status.
We found a significant association between *RUNX3*-SNP rs2236851 and UC (OR 1.61; CI 1.11-2.32, p=0.020). This association was confirmed by the Transmission Distortion Test (p=0.004) and Haplotype Sharing Statistics (p=0.012). Carriership of rs2236851 in UC was associated with pancolitis (OR 1.86; CI 1.08-3.21) and a tendency to association with an early age of onset (OR 1.59; CI 0.98-2.57). For a microsatellite marker in *RUNX3*, an association was found for colonic localization of CD (p=0.022).
For *SLC22A4/5* homozygosity for SNPs rs272893 and rs273900 was significantly associated with CD (OR 2.16; CI 1.21-3.59; p=0.008 and OR 2.40; CI 1.43-4.05 p=0.004 respectively) but not with UC. Both SNPs were associated with an age of onset >40 yrs (p=0.0016 and p=0.0007), ileocolonic localization (p=0.001 and p=0.002), non-stricturing non-penetrating behavior (p=0.020 and p=0.011), and extra-intestinal manifestations (p=0.001 and p<0.001). We did not find any association for the previously described SNPs rs2631367 (-207 G→C) and SNP rs1050152 (1672 C→G) with IBD, CD, UC or any subgroups.
Binary logistic regression analysis revealed an OR of 3.83 (CI 1.26-11.67; p=0.018) for UC patients with homozygosity for one or both associated SNPs rs272893 or rs273900 in *SLC22A4*, and carriership of rs2236851 in *RUNX3*. No evidence for epistasis of *CARD15* with *SLC22A4/5* or for *CARD15* with *RUNX3* was found in IBD, UC or CD or in subsets of UC or CD patients.
Next we analyzed RUNX3 and OCTN1 mRNA expression in inflamed and non-inflamed ileal

and colonic mucosal tissue samples from 30 IBD patients (16 CD and 14 UC) and 6 controls. In controls RUNX3 mRNA expression is evenly distributed throughout the colon and the ileum. OCTN1 expression is higher in the ileum compared to the colon ($p<0.00001$), while the expression level is constant throughout the colon. RUNX3 expression levels are increased in inflamed colonic mucosa compared to non-inflamed colonic mucosa in UC patients ($p=0.01$). OCTN1 expression is decreased in inflamed colonic mucosa compared to non-inflamed colonic mucosa ($p=0.08$).
We conclude that we have provided evidence for the genetic association of *RUNX3* with ulcerative colitis. Furthermore genetic association for Crohn's Disease with *SLC22A4/5* was confirmed, although this involved two other SNPs than previously described. An epistatic effect of *RUNX3* and *SLC22A4* was observed for susceptibility for ulcerative colitis. Our data, combined with the increasingly recognized role of RUNX3 in autoimmunity, suggest an important role for RUNX3 in UC susceptibility.

The next part of the thesis consists of three studies replicating previously described associations in distinctive cohorts using different approaches.
In **chapter 5**, a comparison is made between adult-onset and pediatric-onset CD for genetic variants associated with CD susceptibility. Genetic susceptibility may have a more important role in the aetiology of early-, than of late-onset IBD. After all, early-onset IBD patients were less exposed to environmental factors than late-onset IBD patients. If so, a higher frequency of the gene mutations can be expected in pediatric IBD patients. Next to *CARD15* and *SLC22A4/5,* associations have been found between *DLG5* (Drosophila Discs Large Homologue 5) and the Toll-like receptor 4 (*TLR4)* and IBD. The purpose of this study was to determine the frequency of *CARD15*, *TLR4*, *SLC22A4/5* and *DLG5* mutations in pediatric-onset IBD, to compare these data with adult-onset IBD and to identify genotype-phenotype associations. For *CARD15,* the three known risk-associated variants R702W, G908R and 3020insC and for *TLR4* Asp299Gly and Thr399Ile were determined. For *SLC22A4/5* the 6 SNPs as described in chapter 6 including SNPs -207G→C and 1672C→T were analyzed. Finally, 4 haplotype tagging SNPs in *DLG5* including 113G→A were assessed in 103 pediatric-onset and 696 adult-onset IBD patients. Phenotypic classification was based on disease localization and behaviour. We observed that homozygosity for 3020insC in *CARD15* was more prevalent in pediatric-onset than in adult-onset CD (4.2% v 0.6%, CI 1.2-42.0) and that homozygosity for SNP rs3792876 in *SLC22A4/5* was significantly higher in pediatric-onset CD than in adult-onset CD (6.1% v 1.1%, $p=0.02$). For the other evaluated SNPs no difference in distribution between pediatric-onset and adult-onset IBD could be observed. On a phenotypic level we observed an association of 3020insC and ileal localization for both pediatric-onset and adult-onset groups. Although the numbers in the pediatric cohort were small, we concluded that genetic susceptibility has a more important role in the aetiology of early- than of late-onset CD and that within pediatric-onset CD specific genotype-phenotype associations can be found.
Recently, two novel IBD-associated genes were identified in two large genome-wide case-control analyses. One, an uncommon coding variant (rs11209026) in the gene encoding for the interleukin-23 receptor (*IL23R*) conferred strong protection against CD. It was also shown to be associated with UC in non-Jewish patients. The other, rs2241880 in the autophagy-related

16-like 1 gene (*ATG16L1*) was shown to be associated with CD. Since it is of pivotal importance that genetic associations are confirmed in independent cohorts from different countries, we performed a replication study in **chapter 6** for the two most strongly associated SNPs in *IL23R* and *ATG16L1* in 518 IBD patients (311 CD and 207 UC) and 893 healthy controls. We were also interested in discovering whether these two genes are more generally involved in other common chronic disorders of the gastrointestinal tract and we therefore included a cohort of 508 patients with celiac disease. The case and control samples, as well as 90 CEPH (Centre d'Etude du Polymorphisme Humain) samples included in the International HapMap project, were genotyped by primer extension of multiplex PCR products, followed by a chip-based matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF). We found that the rs11209026 SNP in *IL23R* had a protective effect for IBD in case-control analysis (OR=0.19; 95% CI: 0.10–0.37; p=6.6E-09). This was true for both CD (OR=0.14; CI: 0.06-0.37; p=3.9E-07) and UC (OR=0.33; CI: 0.15-0.73; p=1.4E-03). For *ATG16L1* the rs2241880 SNP was associated with CD susceptibility (OR=1.36; CI: 1.12-1.66; p=0.0017). The population attributable risk for CD is 0.24 for carrying allele G and 0.19 for homozygosity for allele G. No association was found between *IL23R* or *ATG16L1* and celiac disease. Our results confirm the genetic association of *IL23R* with both Crohn's disease and ulcerative colitis, and for *ATG16L1* with Crohn's disease.

It is not only mandatory that previously found associations are confirmed in independent cohorts, these cohorts also need to have sufficient power to detect specific genotype-phenotype interactions. For genetic research in complex genetic disorders as IBD it is important to have large homogenous cohorts of well described and uniformly phenotyped patients. For that reason, on behalf of the Iniative on Crohn and Colitis (ICC), a large nationwide collaborative project was initiated. We performed an association analysis for *DLG5, SLC22A4/5* and *ATG16L1* with IBD, CD, UC and different subsets of CD and UC. Furthermore we evaluated the interaction between these genes. Results are described in **chapter 7.**

DNA samples and phenotypic details of IBD patients from seven University Medical Centers in the Netherlands (University Medical Center Groningen; Academic Medical Center, Amsterdam; VU University Medical Center, Amsterdam; Leiden University Medical Center; Radboud University Nijmegen Medical Center; Erasmus Medical Center, Rotterdam and the University Medical Center Utrecht) were collected. 2937 patients (1696 CD, 1099 UC and 142 with indeterminate colitis) and 1484 healthy controls were included. Phenotypic details were available for 2090 patients (1315 CD / 775 UC). A set of eight SNPs was analyzed for *SLC22A4/5*. This set included, among others, the two known risk-associated SNPs and SNP rs3792876 which was found to be associated to pediatric IBD in chapter 7. Additionally, two markers located at the 3' end and 5' end were chosen to analyze linkage disequilibrium within the IBD5 region. Four SNPs were analyzed for *DLG5*. These SNPs were carefully chosen to identify the four different haplotypes described in the original article. The risk associated SNP rs2241880 was determined for *ATG16L1*.

Firstly, we confirmed the association for CD with the IBD5 haplotype. However, this was not due to specific polymorphisms in the *SLC22A4* and *SLC22A5* genes. Our study was sufficiently powered to detect an association for SNPs rs2631367 (-207G→C) and rs1050152 (1672G→C) in *SLC22A4/5* with an OR of 1.148. However we failed to show any association for these particu-

lar SNPs. We did, however find an association for (subsets of) CD with markers at the 5' and 3' end of the IBD5 locus and with a marker in between *SLC22A4* and *SLC22A5*. This indicates that *SLC22A4/5* are not the CD-associated genes in the IBD5 locus. We further observed an association with an early age of onset and a complicated course of CD with specific SNPs in the region.

Secondly, we did find an association for *DLG5* and IBD, but surprisingly this risk was higher for UC than for CD, in particular with for UC with an extensive phenotype (e.g. leftsided and extensive colitis combined: OR of 1.69). This is a novel finding as none of the previously published studies have found an association with UC for the R30Q variant. Furthermore, we did confirm the association of P1371Q with CD, but could not confirm the protective effect of the haplotype A tagging SNP (rs2289311). However, we did find a protective effect for IBD and CD for the haplotype B tagging SNP. Our results regarding *DLG5* and IBD susceptibility show that there is an association with IBD but this effect is not exclusive for CD. This indicates that *DLG5* is not disease specific but a risk factor for IBD in general.

Furthermore, we could confirm the association of *ATG16L1* with CD. We found a population attributable risk for carrying allele of 0.116. We found the association in particular with CD with stricturing behavior and a history of surgical intervention and ileal or ileocolonic localization..This is the first report of an association of *ATG16L1* with a specific phenotype of CD. We could not find any other associations with UC or any specific subsets, neither on allelic nor on a genotypic level.Most importantly, we show that when combining the information from *ATG16L1*, *DLG5* and the IBD5 haplotype the relative risk for developing CD increases. This risk increases with an increasing number of risk-alleles. This is the first time that this effect is reported for CD. It is consistent with the idea that multiple genes are involved in CD susceptibility, each with a small overall contribution. We demonstrate that it is possible to combine information from multiple common low-penetrance variants to predict the susceptibility for a complex disease. Moreover, an increasing number of risk genotypes or risk alleles was not only associated with CD susceptibility but also with a more complex disease course. Patients with more risk alleles or genotypes had a more severe disease behavior and were more frequently operated. This is an important finding because this emphasizes the fact that it might be possible in the near future to create a genetic risk-profile for a patient to predict the disease course.

The last part of the thesis focuses on azathioprine-toxicity in the treatment of IBD. **Chapter 8** studies the occurrence of side-effects of azathioprine in different diseases. This chapter focuses mainly on acute pancreatitis, which is rarely seen when azathioprine is used for other diseases than Crohn's disease. We performed a retrospective case note survey of 1564 patients using azathioprine after liver or renal transplantation, for systemic lupus erythemathosis, Wegener's granulomatosis, autoimmune hepatitis, rheumatoid arthritis, ulcerative colitis or Crohn's disease. We observed that 11 of 224 patients with Crohn's disease experienced acute pancreatitis (4.9%) compared to 2 / 129 (1.5%) with autoimmune hepatitis, 2 / 388 (0.5%) after renal transplantation, 1 / 254 (0.4%) after liver transplantation and none of the patients using azathioprine for Wegener's granulomatosis (n=85), systemic lupus erythematosis (n=73), Rheumatoid arthritis (n=317) and ulcerative colitis (n=94). The prevalence of azathioprine-induced pancreatitis was higher in patients with Crohn's disease

compared to any other disease ($p<0.05$). Furthermore we observed that azathioprine-toxicity (all toxicity, including acute pancreatitis) necessitating withdrawal occurred significantly more in patients with rheumatoid arthritis (24.6%), ulcerative colitis (21.3%) and Crohn's disease (23.2%) compared to the other patient groups ($p<0,05$).

We conclude that acute pancreatitis is strongly associated with Crohn's disease and rarely occurs with other underlying conditions and that overall azathioprine-induced toxicity and the necessity of withdrawal is more common in inflammatory bowel diseases and rheumatoid arthritis compared to other diseases. The pathogenesis of azathioprine-induced acute pancreatitis is unknown but our analysis suggests that there is a correlation with Crohn's disease, which suggests a disease specific underlying mechanism. Circulating pancreatic antibodies (PAB) have been found in approximately 30 % of patients with CD and were not found in healthy controls, or in patients with other gastrointestinal diseases and in various autoimmune disorders (including autoimmune hepatitis, systemic lupus erythematosis and rheumatoid arthritis). Since PAB and azathioprine-induced pancreatitis are both specific for Crohn's disease, an association or pathogenic role of PAB in azathioprine-induced pancreatitis can be hypothesized. A pathogenic mechanism could be that azathioprine aggravates the inflammation in an already inflamed pancreas in Crohn's disease in PAB positive patients.

In addition to the case-note survey, we evaluated the occurrence of PAB in patients with azathioprine induced pancreatitis in Crohn's disease and in controls with Crohn's disease, hypothesizing that most or all patients with azathioprine-induced pancreatitis have circulating PAB. Two out of eight patients with AZA-AP were positive for PAB (25%), determined by a standardized immunofluorescence method. In the control group of 26 CD patients, two patients (7%) were positive for PAB in serum dilutions of 1:2. The difference in prevalence was not statistically significant. Therefore, although numbers are small, we could not confirm our hypothesis that most or all patients with AZA-AP were PAB positive.

Future perspectives

The future of genetics in IBD research will focus on a number of questions.

1. How will novel IBD associated genes be identified? How will an association with a specific phenotype be firmly established?
2. Once a genetic association is confirmed: how is a particular gene -and the mutations found- involved in IBD susceptibility?
3. How will these findings translate to the treatment of the patient? What will be the role of "pharmacogenomics"?
4. Will genetics be helpful in classifying IBD patients?

1. Future genetic studies in IBD

Numerous case-control studies have been performed concerning the association of a specific candidate gene with IBD. Many of these studies have resulted in positive results but only few of them have been consistently confirmed in independent cohorts. Considerable debate is going on about the appropriateness of applying a candidate-gene approach to complex genetic diseases as IBD. It can be successful in identifying genes with high relative risks as *CARD15*. However, UC and CD and even different subsets of UC and CD have different aetiological mechanisms therefore numerous genes, each with a small overall contribution and a small excessive risk can be involved. For that reason novel IBD associated genes are probably more difficult to identify then *CARD15*, and it is doubtful whether candidate gene studies in small local cohorts will be sufficiently powerful to find new associations.

Next to hypothesis driven candidate gene approaches, genome wide linkage studies have been performed, using markers evenly spaced throughout the genome to identify linkage areas. These studies are obviously very costly. New sets with specific markers throughout the whole genome have been developed and recently the first genome wide case control studies have been performed, identifying two novel IBD associated genes. The gene encoding for the interleukin 23 receptor (*IL23R*) and the autophagy-related 16-like 1 gene were associated with CD and *IL23R* was also associated with UC. To have a homogenous cohort, specific subgroups of IBD -in the case of *IL23R*: CD with ileal localization- were initially studied and subsequently the most informative SNPs were confirmed in independent CD and UC cohorts. We were able to confirm these associations in the Dutch IBD population. Due to statistical deviation, which is invariably associated with this kind of studies, several truly associated genes will not be identified. Therefore, it is still worthwhile to perform a genome wide association study in an independent cohort using these sets of markers. One approach will be to start off with a different subgroup of IBD with a specific unambiguous phenotype like ulcerative pancolitis.

Another important issue is that most genetic research in IBD has been performed in Caucasian patients. Although *CARD15* has been confirmed in many different Caucasian populations, this association could not be found in a Japanese cohort. Probably different genes are involved in IBD susceptibility in different races. It is therefore necessary to generate IBD cohorts in different races to study IBD susceptibility.

To have enough power to detect the association of a specific gene with a particular phenotype, large cohorts need to be studied. These cohorts need to have an adequate uniform description of the clinical characteristics (e.g. the Vienna or Montreal classification). In this thesis we made an effort to establish a nationwide collaborative program for IBD and genetics. On behalf of the Initiative on Crohn and Colitis (ICC) we were able to collect all DNA samples and phenotypic descriptions of patients from seven university medical centres in the Netherlands and set up a preliminary national database. A firm collaboration between all University Medical Centres in the Netherlands presently established for IBD research is.
For future genetic research, whether it is to identify novel susceptibility genes or confirm previously found genes, it is mandatory to study large, well phenotyped, nationwide cohorts and even larger cohorts of matched controls. This requires a specific infrastructure and collaboration between many research centres with sufficient funding and high-throughput molecular platforms.

2. Genetics and Functional studies

Once a firm genetic association is established, the next question is, how a particular gene and the mutations found, are involved in IBD susceptibility. *CARD15* encoding for NOD2 is strongly associated with CD with ileal localization. CD is characterized by an increased activity of NF-κB and NOD2 has been shown to have a role in the activation of NF-κB. However, the precise mechanism, how mutations in *CARD15* and subsequently activation of NF-κB lead to susceptibility of IBD is still only partially understood. Mouse models, in vitro data and studies in humans offer conflicting data. Further studies are necessary and are currently being undertaken to identify how NOD2 variants are involved in IBD susceptibility.
Since mutations in *IL23R* are protective for the development for CD and UC it is tempting to speculate that a defective receptor for IL23 blocks IL23-driven intestinal inflammation. How other associated genes like *DLG5* and *ATG16L1* are involved in IBD susceptibility is almost entirely unknown, and further studies are needed for this matter.
Close collaboration between clinicians, basic scientist involved in IBD research and the geneticists is necessary to proceed in studying functional implications of mutations in IBD associated genes.

3. Genetics and the Treatment of IBD

IBD affect mainly young patients and can cause severe morbidity and have a marked impact on quality of life due to relapsing disease, interference with work and / or studies and so forth but also due to the use of medication and their side-effects. Up till now there is no cure available for IBD and treatment is focussed on the induction of remission and subsequently the maintenance of remission.
Important immunomodulatory drugs, used for the maintenance of remission are the thiopurines, but their use is hampered by the frequent occurrence of side effects in IBD as is shown in this thesis. Polymorphisms in the genes encoding for Thiopurine Methyl Transferase (TPMT) and Inosine Triphosphate Pyrophospatase (ITP-ase) are responsible for a subset of side-effects. Future pharmacogenetic research in IBD will be focussed on identifying patients that are at risk for the development of side effects of thiopurine therapy. In addition phar-

macogenetic research might be able to identify patients that are more likely or less likely to respond to specific therapy (e.g. polymorphisms in glucocorticoid receptors, tumour necrosis factor-α receptors).
Nowadays many so called "biologicals" are being investigated as immunomodulatory medication. They are directed against different components of the inflammatory cascade involved in IBD pathogenesis. This stresses the fact that research in basic immunological and genetic mechanisms of IBD susceptibility is mandatory to clarify these underlying mechanisms. Moreover this highlights also the fact that clinical doctors treating patients have to be aware of these mechanisms, since biologicals will be the most important drugs in the treatment of IBD in the near future.
Although genetic research has not yet led to changes in the clinical management of patients, it will be helpful in selecting patients for a specific therapy and will provide important contributions in the development of new therapeutic approaches.

4. Genetics and Patient classification

Next to genetic associations, an increasing number of serological markers like perinuclear antineutrophil cytoplasmic antibodies (pANCA), antibodies against Saccharomyces cerevisiae (ASCA), and Pancreatic antibodies are known to be associated with different subsets of IBD. Since the successful connection of *CARD15* with ileal CD it is recognized that different phenotypes of CD or UC are characterized by different genetic and serological markers.
For future research, it is therefore of utmost importance that patients are accurately phenotyped according to a well defined clinical classification scheme. The Vienna classification, which is used in the current thesis, is frequently used for genetic studies in CD and includes age of onset (A), disease localisation (L) and diseases behaviour (B). A number of studies have validated this classification; however, several considerations have led to an update of the Vienna classification system during an expert meeting in Montreal in 2005. The main modifications were the introduction of an early age of onset category (< 16 years), the possibility of co-classification of upper gastrointestinal involvement and the inclusion of perianal disease as a disease modifier instead of being a form of penetrating disease. The latter is an important modification because it is recognized that there is no clear association between perianal disease and intra-abdominal penetrating disease.
The aim of rigorous phenotyping is to include serological and genetic markers into the clinical classification system to stratify patients and eventually to predict disease course and response to medical therapy.

Chapter 10

Nederlandse samenvatting en toekomstperspectieven

Samenvatting

De chronische inflammatoire darmziekten (inflammatory bowel diseases - IBD) bestaan uit colitis ulcerosa (CU) en de ziekte van Crohn (Morbus Crohn–MC). Beide zijn chronische ziekten van het maagdarmkanaal die gekenmerkt worden door episodes van actieve ontsteking en episodes van rust. De prevalentie bedraagt 100-200 patiënten per 100.000 inwoners in de westerse wereld. De oorzaak van IBD is maar voor een deel bekend. Er is sprake van een dysregulatie van het immuunsysteem in respons op de normaal aanwezige bacteriën in de darm. Bepaalde omgevingsfactoren zoals roken spelen daar ook een rol bij. Epidemiologische onderzoeken hebben aangetoond dat IBD vaker voorkomt binnen families en bij één- eiige tweelingen is er vaker concordantie van de ziekte dan bij twee-eiige tweelingen. Dit suggereert dat, naast de eerder genoemde factoren, ook een erfelijke factor van belang moet zijn. De ziekte heeft dan ook een multifactoriële achtergrond: de samenstelling van de bacteriën in de darm, de ontregeling van het immuunsysteem, omgevingsfactoren en dus een erfelijke component.
In de afgelopen 10 jaar is er een enorme vooruitgang geboekt in het ontrafelen van de genetische achtergrond van IBD. Op het humane genoom zijn meerdere "risicogebieden" voor het verkrijgen van IBD geïdentificeerd. In 2001 is ontdekt dat mutaties in het gen *CARD15,* dat codeert voor het eiwit NOD2, zijn geassocieerd met het ontstaan van MC. NOD2 is een onderdeel van het aangeboren afweersysteem en herkent specifieke componenten in het membraan van bacteriën. Dit suggereert dat een defect in de herkenning van bacteriën een rol speelt bij het ontstaan van MC. Twee missense mutaties (R702W en G908R) en een frameshift mutatie (L1007fsinsC of 3020insC) zijn geassocieerd met MC met lokalisatie in het terminale ileum.
Omdat er vele fenotypische presentaties zijn van IBD en omdat er meerdere "risicogebieden" zijn geïdentificeerd, worden MC en CU beschouwd als multigene ziekten. Waarschijnlijk zijn meerdere laagpenetrante genen, met elk apart een kleine bijdrage, geassocieerd met (verschillende fenotypes van) MC en CU. Voor genetisch onderzoek naar complexe genetische ziekten als IBD is het erg belangrijk om grote homogene cohorten van welomschreven patiënten te bestuderen. In 2001 is de CODE studie (Chronische Ontsteking van de Darm en Erfelijkheid) gestart in het Universitair Medisch Centrum Groningen. Er is DNA verzameld van IBD patiënten van het Kaukasische ras en hun familieleden. Verder zijn alle patiënten fenotypisch beschreven volgens internationaal geaccepteerde classificatiesystemen zoals de Vienna classificatie voor MC.
Een ander onderzoeksterrein binnen de IBD-genetica is de pharmacogenetica. Er is veel onderzoek gedaan naar het metabolisme van azathioprine. Azathioprine is een purine analoog, dat frequent gebruikt wordt in de behandeling van IBD. Het gebruik van azathioprine wordt echter ernstig beperkt door de bijwerkingen. Polymorfismen in de genen die coderen voor Thiopurinemethyltransferase (TPMT) en inosine triphosphatase (ITP-ase) zijn verantwoordelijk voor een deel van de bijwerkingen van azathioprine.
Dit proefschrift onderzoekt de genetische achtergrond van IBD. Het doel van het proefschrift is om specifieke genotype-fenotype associaties te vinden voor bekende genen en om nieuwe geassocieerde genen te ontdekken. In het eerste deel wordt de associatie voor twee nieuwe

genen met IBD onderzocht. Het tweede deel bestaat uit drie studies waarin eerder gevonden associaties opnieuw worden belicht vanuit verschillende perspectieven. Het laatste deel richt zich op azathioprine therapie en toxiciteit bij IBD.

Hoofdstuk 1 is de introductie van dit proefschrift. De onderzoeksdoelen worden in dit hoofdstuk uiteengezet.

Het focus van dit proefschrift is gericht op specifieke genetische associaties met IBD. Daarom wordt er in **hoofdstuk 2** eerst een uitgebreid overzicht gegeven van de recente literatuur over genetica en IBD. Ondanks veel onderzoek in muizenmodellen, patiënten en *in vitro*, is het nog steeds niet duidelijk hoe mutaties in *CARD15* tot het ontstaan van MC lijden. Verder zijn er vele nieuwe kandidaat-genen ontdekt, waarvan sommige associaties werkelijk bevestigd zijn en er voor andere associaties veel conflicterende studies zijn.Veelbelovende genen zijn *TLR4, MDR1,* NOD1 *(CARD4), DLG5* en het IBD5 locus op chromosoom 5 dat de genen *SLC22A4/5* bevat. Verder wordt er in gegaan op het belang van een uniforme fenotypische classificatie en op de rol van pharmacogenetica in de behandeling van IBD.

Het eerste deel van het proefschrift omvat twee studies die de associatie tussen twee nieuwe kandidaat genen en IBD onderzoeken. In **Hoofdstuk 3** wordt er een case-control onderzoek verricht naar een associatie van het gen dat codeert voor de interleukin-receptor associated kinase-M (*IRAK-M*) en IBD. IRAK-M is een Nuclear Factor-κB (NF-κB) afhankelijke, negatieve regulator van Toll-like receptor signaaltransductie. Toll-like receptoren (TLRs) zijn, net als NOD2, een belangrijk onderdeel van het aangeboren afweersysteem. Eerdere studies toonden dat *IRAK-M* knock-out muizen een verminderde tolerantie hadden voor bacteriën en een verhoogde inflammatoire respons. Gezien de functie van IRAK-M als een negatieve regulator van de TLR-signaaltransductie en de lokalisatie van het gen in een IBD risicogebied op het genoom (IBD2) lijkt het een goede kandidaat voor associatie met IBD. We hebben 542 IBD patiënten (309 MC en 233 CU) en 305 controles bestudeerd. Twee single nucleotide polymorphisms (SNP's) in exonen en 6 microsattelietmerkers werden geëvalueerd en de resultaten werden gestratificeerd voor de *CARD15* status. SNP genotyperingen werden verricht middels Taqman PCR primer / probe sets volgens een gestandaardiseerd protocol. De PCR's voor de microsatteliet markers werden ook volgens een gestandaardiseerd protocol verricht. We vonden geen verschil in distributie van de allelen tussen IBD-patiënten en controles. Wel vonden we dat dragerschap van een van de drie *CARD15* mutaties in combinatie met een microsatteliet marker een verhoogd risico gaf op het ontstaan van CU. Gezien het feit dat er nooit eerder gevonden is dat *CARD15* geassocieerd is met CU en de lage aantallen in deze subgroep is het de vraag of dit een echte associatie is of een vals positief resultaat.
In **hoofdstuk 4** bestuderen we de genetische associatie voor *RUNX3* met IBD. RUNX3 vormt samen met RUNX1 en RUNX2, de runt domain familie van transcriptiefactoren. Deze transcriptiefactoren worden in toenemende mate herkend als belangrijke factoren bij het ontstaan van auto-immuunziekten. Er waren verschillende redenen om *RUNX3* als kandidaat gen te onderzoeken. Ten eerste is het een belangrijke component van de TGF-β signaalroute, welke een belangrijke anti-inflammatoire rol heeft bij IBD. Verder ontwikkelen *Runx3*

knock-out muizen spontaan een colitis. En tenslotte ligt het coderende gen op chromosoom 1p36, een risicogebied voor IBD, op het humane genoom. We bestudeerden 4 SNPs en 4 microsattelietmerkers in en rond het gen *RUNX3* in het CODE cohort. Daarnaast was uit een eerdere publicatie bekend dat de genen *SLC22A4/5* in het IBD5 locus geassocieerd zijn met MC. Een ander artikel toonde dat een SNP die de bindingsplaats voor *RUNX* verstoort in *SLC22A4* geassocieerd is met reumatoïde artritis. Daarom onderzochten we ook zes SNPs in en rond *SLC22A4/5* en de interactie met *RUNX3*. We vonden een significante associatie voor *RUNX3*-SNP rs2236851 met CU (OR1.61; CI 1.11-2.32; p=0.020) en voornamelijk met pancolitis ulcerosa (OR 1.86; CI 1.08-3.21). Deze associatie werd bevestigd door middel van TDT en haplotype sharing statistics. We vonden geen associatie voor de twee eerder beschreven SNPs -207G→C en 1672C→G in *SLC22A4/5*. Wel was er een associatie in geval van homozygotie voor SNPs rs272893 en rs273900 met MC (OR respectievelijk 2.16; CI 1.21-3.59; p=0.008 en OR 2.40; CI 1.43-4.05; p=0.004). Binaire logistische regressieanalyse toonde een OR van 3.83 (CI 1.26-11.67; p=0.018) voor CU wanneer een patiënt drager was van de risico SNP voor *RUNX3* en homozygoot voor een van de twee risico SNP's voor *SLC22A4/5*.
Vervolgens hebben we de expressie van RUNX3 en OCTN mRNA (het eiwit waar *SLC22A4* voor codeert) bepaald in ontstoken en niet ontstoken colon- en ileummucosa van 30 IBD patiënten (16 MC en 14 CU) en 6 gezonde controles. Bij de controles is de expressie van RUNX3 mRNA evenredig verdeeld in het colon en ileum en OCTN1 expressie is in het ileum hoger dan in het colon ($p<0.00001$). RUNX3 mRNA is toegenomen in ontstoken colon in vergelijking met niet ontstoken colon bij de CU patiënten (p=0.01). OCTN1 expressie lijkt lager in ontstoken colon in vergelijking met niet ontstoken colon (p=0.08) bij IBD.
In dit hoofdstuk hebben we aangetoond dat RUNX3 betrokken is bij de pathogenese van CU. Een SNP in een intron van *RUNX3* is genetisch geassocieerd met CU en RUNX3 mRNA is toegenomen in ontstoken colon in vergelijking met niet ontstoken colon. Verder vonden we een associatie voor *SLC22A4/5* met IBD, alhoewel dit andere SNP's betrof dan in eerdere publicaties gevonden werd. Verder lijkt er een epistatisch effect te zijn tussen *RUNX3* en *SLC22A4/5*.

Het volgende deel van het proefschrift bestaat uit drie studies die eerder beschreven genetische associaties voor IBD bestuderen en hiervoor verschillende benaderingen gebruiken.
In **hoofdstuk 5** worden genetische varianten vergeleken tussen een groep volwassen MC patiënten (adult-onset) en een groep patiënten die MC ontwikkelden op de kinderleeftijd (pediatric-onset). Op theoretische gronden kan gesteld worden dat genetische invloeden een belangrijkere rol spelen bij de ontwikkeling van MC op de kinderleeftijd dan bij het ontstaan van MC op de volwassen leeftijd. Pediatric-onset MC patiënten zijn minder lang blootgesteld aan omgevingsfactoren dan adult-onset MC patiënten en derhalve kan verondersteld worden dat MC-geassocieerde genetische varianten vaker voorkomen bij deze groep dan bij adult-onset MC. Naast *CARD15* en *SLC22A4/5* zijn er ook associaties bekend voor *DLG5* (Drosophila Discs Large Homologue 5) en de Toll-like receptor 4 *(TLR4)* en IBD. Het doel van de studie was om de frequentie van *CARD15, TLR4, SLC22A4/5* en *DLG5* varianten in pediatric-onset IBD te bepalen , deze data te vergelijken met adult-onset IBD en specifieke genotype-fenotype associaties te analyseren. De drie bekende MC-geassocieerde mutaties in *CARD15* (R702W, G908R en 3020insC) en Asp299Gly en Thr399Ile voor *TLR4* werden bepaald. Verder werden 6 SNP's in

SLC22A4/5 (inclusief SNPs -207G→C en 1672C→T) en 4 haplotype-markerende SNP's in *DLG5* (inclusief de eerder beschreven geassocieerde SNP 113G→A) bepaald in 103 pediatric-onset en 696 adult-onset IBD patiënten.
Homozygotie voor 3020insC in *CARD15* en homozygotie voor SNP rs3792876 in *SLC22A4/5* kwam significant vaker voor in pediatric-onset dan in adult-onset MC (respectievelijk 4.2% v 0.6%; p= 0.04 en 6.1% v 1.1%; p=0.02). We vonden geen verschil in frequentie tussen beide groepen voor de andere SNPs. We vonden een associatie voor MC met lokalisatie in het ileum voor zowel de pediatric-onset als de adult-onset groep voor3020insC. Alhoewel de getallen in de pediatric-onset groep klein waren, is de conclusie dat de genetische component een belangrijkere rol speelt in het ontstaan van IBD bij pediatric-onset dan bij adult-onset MC.
Recent zijn twee nieuwe MC-geassocieerde genen geïdentificeerd door middel van het gehele genoom overspannende case-control associatiestudies. De één is een zeldzame, coderende variant (rs11209026) in het gen dat codeert voor de interleukine-23 receptor (*IL23R*). Deze variant is geassocieerd met een sterk beschermend effect op het ontstaan van MC. Daarnaast bleek er ook een associatie te zijn met het ontstaan van CU bij niet-joodse patiënten. De andere, rs2241880 in het autophagy-related 16-like 1 gen (*ATG16L1*) was geassocieerd met MC.
Het is van enorm belang dat gevonden genetische associaties worden bevestigd in onafhankelijke cohorten in andere landen. Daarom hebben we in **hoofdstuk 6** een replicatiestudie verricht voor de twee sterkst geassocieerde SNPs in *IL23R* en *ATG16L1* in 518 IBD patiënten (311 MC en 207 CU) en 893 gezonde controles. Daarnaast waren we geïnteresseerd of deze genen ook geassocieerd waren met andere chronische darmziekten. We hebben daarom een cohort van 508 patiënten met coeliakie geïncludeerd. DNA samples van de patiënten en de controles en 90 CEPH (Centre d'Etude du Polymorphisme Humain) samples vanuit het Internationale HapMap project werden gegenotypeerd middels matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF).
SNP rs11209026 SNP in *IL23R* had een beschermend effect voor IBD bij de case-control analyse (OR=0.19; 95% CI: 0.10–0.37; p=6.6E-09). Dit gold zowel voor MC (OR=0.14; CI: 0.06-0.37; p=3.9E-07) als voor CU (OR=0.33; CI: 0.15-0.73; p=1.4E-03). Voor *ATG16L1* was SNP rs2241880 geassocieerd met MC (OR=1.36; CI: 1.12-1.66; p=0.0017). De population attributable risk voor MC is 0.24 wanneer er sprake is van dragerschap van het G-allel en 0.19 wanneer er sprake is van homozygotie is voor het G-allel. We vonden geen associatie voor *IL23R* of *ATG16L1* met coeliakie. Onze resultaten bevestigen de genetische associatie voor *IL23R* met zowel MC als CU en voor *ATG16L1* met MC.
Genetische associaties moeten niet alleen worden bevestigd in onafhankelijke cohorten, deze cohorten moeten ook voldoende power hebben om specifieke genotype-fenotype interacties te kunnen detecteren. Het is belangrijk om grote homogene cohorten met goede uniforme fenotypische beschrijvingen te hebben, om de genetische achtergrond te bestuderen van complexe ziekten als IBD. Er is daarom een groot nationaal samenwerkingsverband opgestart vanuit de Initiative on Crohn and Colitis (ICC). Hierbij zijn al het beschikbare DNA en de fenotypische beschrijvingen van IBD patiënten uit zeven Universitaire Medische Centra in Nederland samengebracht. De resultaten van de associatie analyse voor *DLG5, SLC22A4/5* en *ATG16L1* met IBD, MC, CU en de verschillende fenotypes van MC en CU worden beschreven in **hoofdstuk 7**. Daarnaast wordt de interactie tussen deze genen bepaald.

2937 Patiënten (1696 MC, 1099 CU en 142 met indeterminate colitis) en 1484 gezonde controles werden geïncludeerd. Fenotypische beschrijvingen waren beschikbaar voor 2090 patiënten (1315 MC / 775 CU). Er werden 8 SNP's geanalyseerd voor *SLC22A4/5*. Deze set bevatte onder andere SNPs -207G→C en 1672C→T en SNP rs3792876 welke geassocieerd was met pediatric onset IBD in hoofdstuk 5. Daarnaast werden twee SNPs bepaald gelegen aan het 3' en het 5' einde van het IBD5 locus om linkage disequilibrium met dit haplotype block te bepalen. Er werden 4 SNP's bepaald voor *DLG5* die dusdanig waren gekozen dat alle 4 haplotypes uit het originele artikel konden worden geïdentificeerd. Verder werd SNP rs2241880 bepaald voor *ATG16L1*.
Allereerst vonden we een associatie voor MC met het IBD5 haplotype. We vonden geen associatie met het de twee eerder beschreven SNPs 207G→C en 1672C→T. Wel vonden we dat *SLC22A4/5* geassocieerd zijn met MC maar dat dit niet onafhankelijk was van het IBD5 locus. Er was sprake van een sterk linkage disequilibrium tussen alle SNPs in en rond *SLC22A4/5*. *SLC22A4/5* kunnen dus wel de MC-geassocieerde genen zijn in het IBD5 locus, maar het is net zo waarschijnlijk dat andere genen in dit gebied de werkelijke MC-geassocieerde genen zijn. Verder vonden we dat het IBD5 locus geassocieerd is met het ontstaan van MC op jonge leeftijd en met een gecompliceerd ziektebeloop.
Ten tweede vonden we een associatie voor de R30Q (rs1248696) variant in *DLG5* en IBD. Verrassend genoeg was de associatie sterker voor CU dan voor MC. Dit is een nieuwe bevinding, want in eerder publicaties worden er alleen associaties met MC beschreven. Verder konden we de associatie voor P1371Q (rs2289310) met MC bevestigen, maar vonden we geen beschermend effect voor de haplotype A markerende SNP.
Ten derde konden we ook in dit cohort de associatie voor *ATG16L1* met MC bevestigen. De associatie werd met name gevonden voor MC met lokalisatie in het terminale ileum, met stricturerende ziekte en met chirurgische interventie. We vonden geen associatie met CU of met andere fenotypes.
Vervolgens hebben we de interactie tussen de genen geanalyseerd, door gebruik te maken van de sterkst met MC-geassocieerde SNP's in elk gen. We konden geen specifieke interactie vinden tussen de genen. MC patiënten hadden echter wel significant meer risicoallelen dan controles. Vervolgens konden we aantonen dat het risico op het ontstaan van MC toeneemt met het aantal risicoallelen. Dit is voor het eerst dat dit aangetoond is voor MC en past in het concept dat er meerder genen betrokken zijn bij het ontstaan van MC, die elk onafhankelijk bijdragen aan het ontstaan van de ziekte. Verder hebben we middels ordinale regressie analyse aangetoond dat een gecompliceerd beloop van de ziekte (stricturerende of penetrerende ziekte en chirurgische interventie) geassocieerd was met een toenemend aantal risicoallelen. Dit is een belangrijke bevinding omdat het in de toekomst mogelijk moet zijn om een genetisch risicoprofiel van een MC patiënt te maken om het verloop van de ziekte te voorspellen.

Tenslotte bestuderen we in **hoofdstuk 8** de incidentie van bijwerkingen van azathioprine in patiënten met verschillenden ziekten. De studie is in het bijzonder gericht op het ontstaan van azathioprine- geïnduceerde pancreatitis, wat voorkomt bij de behandeling van MC, maar zelden of nooit wanneer azathioprine gebruikt wordt voor andere ziekten. We hebben een retrospectieve studie verricht van 1564 patiënten die azathioprine gebruikten na een lever

of een niertransplantatie, voor systemische lupus erythematosis, Wegener's granulomatosis, autoimmuun hepatitis, reumatoïde arthritis, CU of MC. We vonden dat acute pancreatitis ten gevolge van azathioprine bij 11 van de 224 (4.9%) patiënten met MC ontstond. Dit is significant ($p<0.05$) vaker dan bij autoimmuun hepatitis (2 / 129 - 1.5%), na niertransplantatie (2 / 388 – 0.5%), na levertransplantatie (1 / 254 – 0.4%), bij systemische lupus erythematosis (0 / 73), Wegener's granulomatosis (0 / 85) , reumatoïde artritis (0 / 317) en CU (0 / 94). Verder moesten patiënten met MC (23.2%), CU (21.3%) en reumatoïde artritis (24.6%) vaker stoppen met azathioprine ten gevolge van alle bijwerkingen dan de andere patiënten groepen ($p<0.05$).
De conclusie is dat azathioprine-geïnduceerde acute pancreatitis sterk geassocieerd is met MC en zelden ontstaat wanneer azathioprine gebruikt wordt voor andere ziekten en dat azathioprine vaker gestopt wordt in verband met de bijwerkingen bij IBD en reumatoïde artritis dan bij andere ziekten.
De oorzaak van azathioprine geïnduceerde acute pancreatitis is onbekend.Het feit dat wij een sterke associatie vonden met MC suggereert een ziekte specifiek mechanisme. Circulerende antilichamen tegen het pancreas (Pancreatic antibodies – PAB) worden in ongeveer 30% van alle MC patiënten gevonden en niet bij gezonde controles of patiënten met andere gastrointestinale aandoeningen. Aangezien zowel PAB als azathioprine geïnduceerde acute pancreatitis specifiek zijn voor MC, kan een associatie of pathogene rol worden verondersteld. Een mogelijkheid zou kunnen zijn dat azathioprine een acute verergering uitlokt van een al enigszins ontstoken pancreas bij PAB positieve MC patiënten. We onderzochten nog de aanwezigheid van PAB in acht MC patiënten met een azathioprine-geïnduceerde pancreatitis en in 26 controles met MC. We vonden geen verschil in prevalentie in beide groepen en konden derhalve onze hypothese niet bevestigen. Wel vonden we een lagere frequentie (12%) van PAB in vergelijking met eerdere studies bij MC.

Toekomstperspectieven

Er zijn verschillende vragen waarop het onderzoek naar de genetische achtergrond van IBD zich in de toekomst zal richten.

1. Op welke manier zullen nieuwe genetische associaties met IBD worden vast gesteld? Hoe zullen associaties met specifieke fenotypes worden gevonden?
2. Wanneer een genetische associatie gevonden en bevestigd is: Hoe is dit gen – en de gevonden mutaties - betrokken bij het ontstaan en het beloop van IBD?
3. Hoe zullen deze bevindingen worden vertaald naar de behandeling van de patiënt? Wat zal de rol zijn van "pharmacogenetics" bij de behandeling van IBD?
4. Zal de genetische achtergrond een rol spelen in het classificeren van IBD patiënten?

1. Toekomstige studies naar nieuwe IBD-geassocieerde genen

In de afgelopen jaren zijn er zeer veel kandidaat-gen studies verricht, waarin de associatie met een specifiek gen met IBD werd onderzocht. Veel van deze studies hebben positieve resultaten opgeleverd, maar er zijn maar weinig van deze associaties bevestigd in onafhankelijke cohorten. Derhalve is er ook veel discussie over de waarde van kandidaat-gen studies in complexe ziekten als IBD. Genen met een hoog relatief risico zoals *CARD15* zullen worden geïdentificeerd. Echter, niet alleen tussen MC en UC, maar vrijwel zeker ook binnen de twee aandoeningen zijn er belangrijke verschillen in de pathogenese. Dit komt waarschijnlijk omdat meerdere genen elk met een lage penetrantie en kleine bijdrage aan de pathogenese betrokken zijn bij het ontstaan van verschillende fenotypes van IBD. Nieuwe genetische associaties zullen dan ook moeilijker te vinden zijn dan *CARD15* en het is dan ook te betwijfelen of kandidaat-gen studies in kleine lokale cohorten genoeg power hebben om deze te vinden.

Recent zijn de resultaten van de eerste het gehele genoom omvattende case-control analyses bij IBD gedaan waarbij 100.000 tot 500.000 geselecteerde SNP's verspreid over het genoom werden geanalyseerd. Deze SNP's werden bepaald in een homogeen cohort IBD patiënten (in het geval van *IL23R* bij MC patiënten met lokalisatie in het terminale ileum). De gevonden associaties werden vervolgens bevestigd in onafhankelijke MC en CU cohorten. Ook wij konden deze associaties bevestigen in ons cohort. Ten gevolge van de enorme hoeveelheid SNP's en de daaraan inherente statistische deviaties, zullen er altijd nog werkelijk geassocieerde genen onontdekt blijven met deze benadering. Het is daarom zeker nog zinvol om dergelijke genoom omspannende associatie-analyses te verrichten in onafhankelijke cohorten. Dit zou bijvoorbeeld gedaan kunnen worden in een andere subgroep van IBD met een welomschreven fenotype, zoals pancolitis ulcerosa.

Een ander belangrijk aspect is dat alle gevonden genetische associaties bestudeerd zijn bij IBD patiënten van het Kaukasische ras. *CARD15* blijkt bijvoorbeeld niet geassocieerd te zijn met IBD in de Aziatische populatie. Waarschijnlijk zijn bij verschillende rassen andere genen betrokken bij het ontstaan van IBD. In de toekomst zullen er dan ook goede studies moeten worden verricht in cohorten IBD patiënten van verschillende rassen.

Om specifieke genotype-fenotype associaties te vinden moeten grote cohorten worden bestudeerd. Daarom hebben we binnen de Initiative on Crohn and Colitis (ICC), een landelijk samenwerkingsverband tussen alle Universitair Medische Centra in Nederland, een project opgestart op het gebied van IBD-genetica. We hebben hierbij alle DNA samples van zeven universitair medische centra verzameld en een nationale DNA database opgezet. Van een groot deel van deze patiënten zijn adequate uniforme fenotypische beschrijvingen beschikbaar. Voor studies naar de genetische achtergronden van IBD zijn grote samenwerkingsverbanden noodzakelijk met specifieke infrastructuren, voldoende financiële ondersteuning en zogenaamde "high throughput" platformen die grote hoeveelheden SNPs kunnen analyseren.

2. Genetica en functionele studies

Wanneer een genetische associatie bij herhaling bevestigd is in onafhankelijke cohorten, is de volgende vraag hoe een specifiek gen en de gevonden mutaties betrokken zijn bij de pathogenese van IBD. In het geval van *CARD15* zijn er, ondanks veel onderzoek, tegenstrijdige bevindingen in de literatuur. De vraag of er sprake is van een toename of een afname van de functie van NOD2 is tot op heden maar gedeeltelijk beantwoord. Voor bijvoorbeeld mutaties in *IL23-R,* die een beschermend effect hebben op het ontstaan van IBD, is het te verwachten dat een defecte IL23 receptor de IL23 gedreven intestinale ontsteking onderdrukt. Op dit moment worden er verschillende studies naar dit onderwerp gedaan. Voor andere geassocieerde genen zoals *ATG16L1* en *DLG5* is er nauwelijks informatie beschikbaar over hun rol in de pathogenese van IBD. Basale wetenschappers, genetici en clinici zullen nauw moeten samenwerken om de functionele implicaties van gevonden genetische associaties te bestuderen.

3. Genetica en de behandeling van IBD

IBD komt voornamelijk voor bij jonge mensen en kan tot ernstige morbiditeit leiden. De ziekte kan grote invloed hebben op de kwaliteit van leven ten gevolge van de ziekteactiviteit maar ook ten gevolge van operaties en bijwerkingen van medicijnen.

Er zijn genetische polymorfismen bekend die verantwoordelijk zijn voor bijvoorbeeld een deel van de bijwerkingen van azathioprine. Pharmacogenetisch onderzoek zal er op gericht zijn om patiënten te identificeren die wel of geen bijwerkingen op specifieke medicamenten zullen ontwikkelen. Daarnaast zal het in de toekomst misschien ook mogelijk zijn om patiënten te identificeren die in meer of mindere mate zullen responderen op specifieke medicatie (bijvoorbeeld door het bepalen van polymorfismen in glucocorticoïd receptoren of tumor necrosis-α receptoren.)

Tegenwoordig worden er veel zogenaamde "biologicals" onderzocht voor de behandeling van IBD. De meeste van deze "biologicals" zijn gehumaniseerde antilichamen of fragmenten daarvan. Deze antilichamen zijn gericht tegen specifieke componenten van de inflammatoire cascade van IBD. Dit onderstreept nogmaals het feit dat basaal immunologisch en genetisch onderzoek noodzakelijk is om deze inflammatoire mechanismen te doorgronden. Daarnaast dient ook de maag-darm-leverarts die patiënten behandelt, kennis te hebben van deze mechanismen, omdat de "biologicals" waarschijnlijk de belangrijkste componenten zullen zijn voor de behandeling van IBD in de nabije toekomst.

4. Genetica en de classificatie van patiënten

Voor wetenschappelijk onderzoek naar IBD, dienen patiënten uniform en volgens internationaal geaccepteerde classificatiesystemen te worden beschreven. Recent is er een update verschenen van de Vienna classificatie: de Montreal classificatie die goed bruikbaar lijkt voor zowel MC als CU. Naast genetische associaties zijn er ook in toenemende mate circulerende antilichamen (Antineutrophil cytoplasmic antibodies - ANCA, antibodies against Saccharomyces cerevisiae – ASCA etc) geïdentificeerd die met specifieke fenotypes van IBD zijn geassocieerd. Het doel is om tot een classificatieschema te komen waarin de klinische (fenotypische) beschrijving, de circulerende antilichamen en de genetische polymorfismen in kunnen worden geïncorporeerd. Dit kan clinici helpen om het toekomstige ziekteverloop van een patiënt te voorspellen en het beleid daarop aan te passen.

List of publications

Weersma RK, Peters FT, Oostenbrug LE, van den Berg AP, van Haastert M, Ploeg RJ, Posthumus MD, Homan van der Heide JJ, Jansen PL, van Dullemen HM. Increased incidence of azathioprine-induced pancreatitis in Crohn's disease compared with other diseases. Aliment Pharmacol Ther. 2004;20:843-50

Weersma RK, van Dullemen HM, Kleibeuker JH, Ploeg RJ, Dijkstra G. Treatment of severe ulcerative colitis. Ned Tijdschr Geneeskd. 2006;150:12-7.

Weersma RK, Thijs WJ, Vosmaer GD, Koornstra JJ. Capsule endoscopy in Klippel - Trenaunay syndrome. Endoscopy. 2007 Feb. 26 DOI 10.1055:/s-2006-945129.

Weersma RK, Limburg AJ, Karrenbeld A, Koornstra JJ. Editor's quiz: iron deficiency anaemia 10 years after small bowel resection in infancy. Gut. 2007;56: 463,488.

Weersma RK, Oostenbrug LE, Nolte IM, Van Der Steege G, Oosterom E, Van Dullemen HM, Kleibeuker JH, Dijkstra G. Association of interleukin-1 receptor-associated kinase M (IRAK-M) and inflammatory bowel diseases. Scand J Gastroenterol. 2007;42:827-33.

de Ridder L, **Weersma RK**, Dijkstra G, van der Steege G, Benninga MA, Nolte IM, Taminiau JA, Hommes DW, Stokkers PC. Genetic susceptibility has a more important role in pediatric-onset Crohn's disease than in adult-onset Crohn's disease. Inflamm Bowel Dis. 2007;13:1083-92.

Weersma RK, van Dullemen HM, van der Steege G, Nolte IM, Kleibeuker JH, Dijkstra G. Inflammatory bowel diseases and genetics: current affairs. Aliment Pharmacol Ther. *Accepted for publication*

Weersma RK, Zhernakova A, Nolte IM, Lefebvre C, Rioux JD, Mulder F, van Dullemen HM, Kleibeuker JH, Wijmenga C, Dijkstra G. ATG16L1 and IL23R are associated with inflammatory bowel diseases but not with celiac disease in the Netherlands. Am J Gastroenterol. *Accepted for publication*

van der Wouden EJ**, Weersma RK.** Endoscopy for obstructive jaundice. Neth J Med, *Accepted for publication*

Weersma RK, Batstra MR, Kleibeuker JH, van Dullemen HM. Are pancreatic autoantibodies a risk factor for azathioprine induced pancreatitis in Crohn's disease? *Submitted*

Weersma RK, Zhou L, Nolte IM, van der Steege G, van Dullemen HM, Oosterom E, Bok L, Peppelenbosch MP, Faber KN , Kleibeuker JH , Dijkstra G. The Runt-Related Transcription Factor 3 is associated with ulcerative colitis and shows epistasis with Solute Carrier Family 22, members 4 and 5. *Submitted*

Weersma RK, Stokkers PCF, van Bodegraven AA, van Hogezand RA, Verspaget HW, de Jong DJ, van der Woude CJ, Oldenburg, Linskens RK, van der Steege G, Hommes DW, Crusius JB, Wijmenga C, Nolte IM, Dijkstra G. Molecular prediction of disease risk and severity in a large Dutch Crohn's disease cohort. *Submitted*

Wapenaar MC, Monsuur AJ, van Bodegraven AA, **Weersma RK**, Bevova MR, Howdle P, Holmes G, Mulder CJ, Dijkstra G, van Heel DA, Wijmenga C. Association of tight junction genes UMCU1 and UMCU2 with gluten sensitive enteropathy and inflammatory bowel disease implies a common barrier defect**.** *Submitted*

Kaser A , Lee AH, Lefebvre C , Glickman JN, Tilg H, Nieuwenhuis EE, Annese V, Brant SR, Cho J, Duerr RH, van der Velden AWM, Starnbach MN, Silverberg M, Taylor KD, **Weersma RK**, Wijmenga C, Higgins DE, Rioux JD, Glimcher LH, Blumberg RS. Coordinate regulation of Paneth cell function and intestinal inflammation by XBP-1 as a risk factor for human IBD. *Submitted*

Weersma RK, de Ruiter AJ, Peters FTM. Peroral cholangioscopic treatment of complex common bile duct calculi by Holmium: yttrium aluminum garnet laser. *Submitted*

De Graaf APJ, **Weersma RK**, Thijs WJ, Limburg AJ, Koornstra JJ. Impact of capsule endoscopy on patient management: a 1-year follow-up study. *Submitted*

Van der Heide F, Dijkstra A, **Weersma RK**, Albersnagel F, van der Logt EMJ, Faber KN, Sluiter WJ, Kleibeuker JH, Dijkstra G. Different effects of active and passive smoking on the disease course of ulcerative colitis and Crohn's disease. *Submitted*

Curriculum Vitae

Rinse Weersma werd op 5 oktober 1972 geboren te Delfzijl. In 1991 behaalde hij zijn VWO-diploma aan het Ommelander College te Appingedam. In september van dat jaar startte hij de studie geneeskunde aan de Rijksuniversiteit Groningen. Tijdens zijn studie werd er veel gesport, gemusiceerd en gereisd. Het doctoraal werd behaald in 1996 en na zijn co-schappen in het Martini Ziekenhuis te Groningen, legde hij het artsexamen af in 1998. Vervolgens werkte hij een jaar als assistent interne geneeskunde in het Deventer Ziekenhuis. Na een paar maanden reizen door Azië, startte hij in april 2000 met de opleiding tot internist in het Martini Ziekenhuis te Groningen (opleider Dr. J.D.M. Gökemeyer). Gedurende deze jaren werd zijn enthousiasme voor de maag-, darm- en leverziekten (MDL) gewekt en startte hij in juni 2003 met zijn vervolgopleiding tot MDL-arts in het Medisch Spectrum Twente te Enschede (opleider Dr. J.J.Kolkman). In juni 2004 vervolgde hij zijn opleiding in het Universitair Medisch Centrum Groningen (opleider Prof. dr. J.H.Kleibeuker). Aldaar werd zijn interesse gewekt voor de wetenschap en begon hij met het onderzoek naar de genetische achtergrond van inflammatoire darmziekten zoals dat beschreven is in dit proefschrift. Sinds juni 2006 is hij werkzaam als MDL-arts in het Universitair Medisch Centrum Groningen met als speciale aandachtsgebieden inflammatoire darmziekten en geavanceerde endoscopieën.

Rinse woont samen met Wencke Veenstra en zij hebben één dochter, Ilpha.

Dankwoord

Promoveren naast een fulltime opleiding en nu naast mijn werkzaamheden als MDL-arts in het UMCG, was onmogelijk geweest zonder de hulp van velen.

Allereerst dank aan alle patiënten en familieleden die belangeloos bloed hebben afgestaan voor het wetenschappelijk onderzoek. Daarnaast ben ik ook veel dank verschuldigd aan Liekele Oostenbrug die op geweldige wijze het genetisch onderzoek in Groningen heeft opgestart.

Prof. Dr. J.H. Kleibeuker, beste Jan. Dank voor je perfecte en rustige begeleiding bij de totstandkoming van dit proefschrift en je grote rol tijdens de laatste twee jaren van mijn opleiding. Het vertrouwen en de ruimte die je mij geeft in mijn ontwikkeling als MDL-arts en wetenschapper waardeer ik enorm.

Prof. Dr. C. Wijmenga, beste Cisca, je bent pas in een latere fase betrokken geraakt bij mijn promotie, maar je had een grote rol bij de laatste artikelen. Ik ben erg blij met je komst naar Groningen en heb er alle vertrouwen in dat we nog veel projecten samen tot een goed einde zullen brengen.

Dr. G. Dijkstra, beste Gerard. Wat moet ik hier nu zeggen? We spreken elkaar dagelijks en hebben het vaak over het vak, de wetenschap, maar ook over alles wat daar bij hoort. Ik heb groot respect voor je enthousiasme, je kennis maar ook voor je integriteit. Jij was de stuwende kracht achter dit proefschrift, waarvoor grote dank.

Dr. H.M. van Dullemen, beste Hendrik. Ook jou spreek ik dagelijks en alles wat gezegd moet worden, hebben we terloops al besproken. Ook jou ben ik veel dank verschuldigd. Niet alleen voor het opstarten van het genetisch onderzoek bij IBD in het UMCG en je begeleiding, maar in het bijzonder ook voor de grote rol die je hebt gespeeld in mijn opleiding tot 'advanced endoscopist.'

De leden van de leescommissie, Prof. Dr. D.W.Hommes, Prof. Dr. R.M.W. Hofstra en Prof. Dr. R.J.Ploeg wil ik bedanken voor hun snelle en kritische beoordeling van dit proefschrift.

De voormalige genotyperingsunit bestaande uit Dr. G. van der Steege, Dr. I.M. Nolte en E. Oosterom is tegenwoordig helaas niet meer rechtstreeks betrokken bij het IBD onderzoek, maar is van cruciaal belang geweest voor dit proefschrift.
Beste Gerrit, bedankt voor je begeleiding en de tijd die je genomen hebt om mij, tussen het bespreken van de alledaagse beslommeringen door, de beginselen van het genetisch onderzoek bij te brengen.
Beste Ilja, zonder jou was dit proefschrift nooit afgekomen. Ik ben nog steeds onder de indruk hoe jij uit gigantische datasets, met het grootste gemak de juiste berekeningen tovert. Hoewel je niet meer bij de afdeling genetica werkt en je je hebt verstopt in een uithoek van

het ziekenhuis, ben je nog steeds enthousiast aan het rekenen aan onze IBD data. Ik hoop dan ook dat onze samenwerking nog lang mag duren.
Beste Elvira, de hoeveelheid werk die jij, altijd goedgemutst, verricht hebt voor dit proefschrift is enorm. In het bijzonder de inspanning die je hebt geleverd om het DNA van het landelijke cohort op tijd te genotyperen heb ik erg gewaardeerd.

Dank ook aan alle medeonderzoekers en co-auteurs. Met name wil ik de leden van de 'Initiative on Crohn and Colitis' voor het in mij gestelde vertrouwen, Lissy de Ridder voor de prettige samenwerking en Klaas Nico Faber voor de prima samenwerking met het lab, bedanken.

Dan de overige stafleden van de afdeling Maag Darm Leverziekten in het UMCG; Els Haagsma, Aad van de Berg, Bram Limburg, en Frans Peters. Eerst mijn opleiders en nu mijn collega's; bedankt voor jullie begeleiding tijdens mijn opleiding tot MDL-arts. Uiteraard ook dank aan het "jongste" staflid en mijn kamergenoot Jan Jacob Koornstra. JJ: we gaan als enthousiaste jonge honden stug verder met "Groningen op de kaart zetten".

Mijn directe collega's tijdens mijn opleiding tot MDL-arts Pascale Dekkers, Louktje Wormmeester, Antoine Flierman en Willem Thijs, bedankt voor de goede tijd en collegialiteit. Willem, ik hoop nog vele congressen met je te bezoeken.

Speciale dank voor de prima samenwerking en goede sfeer aan alle verpleegkundigen en andere medewerkers van het endoscopiecentrum en afdeling E3VA en aan Gonny Thijn, Petra Wetterauw, Bert Boer en Monique Chedalal-Bhawan van de MDL secretariaten.

Mijn opleider interne geneeskunde Dr. JD Gökemeyer, een groot clinicus en een voorbeeld voor vele artsen in opleiding, en de maatschap interne geneeskunde wil ik bedanken voor de goede jaren in het Martini Ziekenhuis Groningen.

Mijn opleider Dr. J.J. Kolkman in het Medisch Spectrum Twente te Enschede. Beste Jeroen, jouw onnavolgbare enthousiasme en de ernst waarmee je je taak als opleider opvat, zijn geweldig. Ook de rest van de MDL staf in Enschede en toenmalig collega MDL arts i.o. Peter Mensink ben ik dankbaar. Mijn jaar als onbestorven weduwnaar in Enschede werd een stuk draaglijker door de goede sfeer.

Dan verlaten we nu het ziekenhuis en komen we bij de mensen die het dichtst bij me staan.

Mijn paranimfen Otto Maarsingh en Jeroen van Zanten. Ot, vanaf het eerste jaar geneeskunde en onze reis door India en Nepal zijn we onafscheidelijk. Onze vriendschap is onvoorwaardelijk en oneindig. Ik ben verschrikkelijk blij dat je naast me staat op deze dag. Joene, het is ongelooflijk, maar onze vriendschap bestaat al vanaf de kleuterschool en sindsdien delen we belangrijke periodes in ons leven. Ik ben trots dat een groot academicus als jij naast me staat vandaag.

Dan heel, heel, heel veel dank aan alle familie en lieve vrienden die mij altijd weer herinneren aan de echt belangrijke zaken in het leven: Martijn en Famke, Auk en Hammy, Wytze en Steffie, Tom en Chantal, Erik en Trudy, Froukje, Chris en Ageeth, Rolf en Berthilde, Klaas en Monique, de HHS crew, de manluu van de Bond tegen Harries en alle andere vrienden die mij hebben bijgestaan in de afgelopen jaren.

Mijn ouders, Bauke en Agnes, zonder jullie nimmer aflatende steun, liefde, vertrouwen en oneindige trots was ik niet de persoon geworden die ik nu ben. Dank voor alles.

Ook al worden er veel mensen genoemd in bovenstaande regels en is zo'n promotie best leuk, uiteindelijk draait alles maar om twee vrouwen: Wencke en Ilpha. Lieve Wencke, ons leven samen is zo geweldig, dat we soms denken dat het allemaal niet echt is en we het zelf verzinnen. Alles is met de komst van Ilpha alleen nog maar mooier geworden. Het is heerlijk om mijn leven met jullie te delen.

Rinse Weersma